

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 8.04852 Seconds  
(without alignments)  
418.411 Million cell updates/sec

Title: US-10-792-311-1  
Perfect score: 166  
Sequence: 1 AKKYAKKEKAAYKKAKEAAAEAAAEAYEA 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84	50.6	384	B43592	outer membrane prote
2	80.5	48.5	461	T03561	hypothetical prote
3	76	45.8	347	B83525	TolA protein PA097
4	74	44.6	372	G64064	outer membrane int
5	72	43.4	210	A25550	histone H1 - sea u
6	71.5	43.1	211	A28100	histone H1-beta, e
7	70.5	42.5	445	T50972	probable zuotin (i
8	70	42.2	206	S09388	histone H1 - sea u
9	70	42.2	277	T34625	probable NLP/p60 f
10	70	42.2	1015	JC6552	DNA topoisomerase
11	70	42.2	1130	T34081	hypothetical prote
12	69.5	41.9	421	JV0057	TolA protein - Esc
13	67	40.4	394	F90725	membrane spanning
14	67	40.4	394	G85576	membrane spanning
15	67	40.4	1021	G75403	DNA topoisomerase
16	66.5	40.1	217	A26721	histone H1-gamma,
17	66.5	40.1	1052	A44937	kinetoplast-associ
18	66	39.8	231	S59589	histone H1 - Chlam
19	66	39.8	318	H96535	hypothetical prote
20	66	39.8	1390	S51364	sperm tail-specifi
21	65.5	39.5	214	G70673	probable hupB - My
22	65.5	39.5	1528	A60338	surface antigen A
23	65	39.2	101	H59099	hypothetical prote
24	64.5	38.9	185	A32137	histone H1-delta -
25	64	38.6	241	JN0748	histone H1-II - Vo
26	64	38.6	335	T21503	hypothetical prote
27	63.5	38.3	621	A57591	Id-associated prot
28	63	38.0	248	1 HSRU1P	histone H1, gonada
29	63	38.0	261	JN0747	histone H1-I - Vol

30 63 38.0 683 2 H96835  
31 63 38.0 735 2 T49622  
32 63 38.0 1701 2 T09127  
33 62.5 37.7 182 2 S61926  
34 62.5 37.7 288 2 T06257  
35 62.5 37.7 826 2 A01683  
36 62 37.3 206 1 HSTR1P  
37 62 37.3 244 2 S40436  
38 62 37.3 266 2 T10644  
39 62 37.3 376 2 AG0592  
40 62 37.3 442 2 T39683  
41 62 37.3 628 2 T08942  
42 62 37.3 827 2 A01311  
43 62 37.3 898 2 A82298  
44 62 37.3 899 2 B87553  
45 61.5 37.0 217 2 S29309

ALIGNMENTS

RESULT 1

B43592

outer membrane protein TmpB - Treponema phagedenis

C:Species: Treponema phagedenis

C>Date: 30-Jan-1993 #sequence\_revision 12-Mar-1993 #text\_change 09-Jul-2004

C:Accession: B43592

R:Yelton, D.B.; Limberger, R.J.; Curci, K.; Malinosky-Rummell, F.; Slivienksi, L.; Schoultz, J.; Infect. Immun. 59, 3685-3693, 1991

A:Title: Treponema phagedenis encodes and expresses homologs of the Treponema pallidum Tm

A:Reference number: A43592; MUID:91372983; PMID:1894368

A:Accession: B43592

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-384 <YEL>

A:Cross-references: UNIPROT:P29720; GB:M58563; NID:g155066; PIDN:AAA27480.1; PID:g155067

A:Note: the authors translated the codon TTC for residue 316 as Tyr, and CGA for residue

C:Keywords: membrane protein

Query Match 50.6%; Score 84; DB 2; Length 384;  
Best Local Similarity 64.9%; Pred. No. 0.092;  
Matches 24; Conservative 5; Mismatches 6; Indels 2; Gaps 2;

QY 1 AKKYAKKEKAAYKKAKEAAAEAAAEAYEA 35  
DB 218 AREMAAKEKAAYKKAKEAAAEAAAEAYEA 254

RESULT 2

T03561

hypothetical protein - Rhodobacter capsulatus

C:Species: Rhodobacter capsulatus

C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004

C:Accession: T03561

R:Vlicek, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M. Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997

A:Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus SB1003.

A:Reference number: Z14955; MUID:97404404; PMID:9256491

A:Accession: T03561

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-461 <VLCS>

A:Cross-references: UNIPROT:O68124; EMBL:AF010496; NID:g3128256; PIDN:AAC16214.1; PID:g31

C:Genetics:

A:Map position: 1

Query Match 48.5%; Score 80.5; DB 2; Length 461;  
Best Local Similarity 60.0%; Pred. No. 0.25;  
Matches 21; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 1 AKKYAKKEKAAYKKAKEAAAEAAAEAYEA 35  
DB 30 AKEVVEKAAYKKAKEAAAEAAAEAYEA 63

hypothetical prote  
hypothetical prote  
probable erythrocy  
histone H1 homolog  
histone H1 (clone  
penicillin-binding  
histone H1 - rainb  
histone H1 - midge  
hypothetical prote  
TolA protein limpo  
zuotin-like protei  
proton pump intera  
penicillin-binding  
translation initia  
DNA topoisomerase  
hypothetical prote



```
Query Match      42.5%; Score 70.5; DB 2; Length 445;
Best Local Similarity 57.1%; Pred. No. 2.6;
Matches 20; Conservative 4; Mismatches 10; Indels 1; Gaps 1;

QY 2 KKYAKKEKAAYKKEAKAKAA-EAAAKAEAYEA 35
   |||::|||::|||::|||::|||::|||::|||
Db 304 KKRLEKEAAEKAAEBAEAAKAAEAAEAAKEBEKA 338

RESULT 8
S09388
histone H1 - sea urchin (Parechinus angulosus)
C:Species: Parechinus angulosus (angulate urchin)
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S09388
R:Hilli, C.S.; Martin, S.R.; Thomas, J.O.
EMBO J. 8, 2591-2599, 1989
A>Title: A stable alpha-helical element in the carboxy-terminal domain of free and chromatin-associated histone H1
A:Reference number: S09388; MUID:90060019; PMID:2583125
A:Accession: S09388
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-206 <HIL>
A:Cross-references: UNIPROT:Q7M409
C:Superfamily: histone H1
C:Keywords: chromosomal protein

Query Match      42.2%; Score 70; DB 2; Length 206;
Best Local Similarity 56.7%; Pred. No. 1.6;
Matches 17; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 3 KYAKKEKAAYKKEAKAKAAEAAAKAEAA 32
   |||::|||::|||::|||::|||::|||
Db 114 KKAKTSAAKAKKAKARARAAAKAKAAA 143

RESULT 9
T34625
probable NLP/P60 family secreted protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T34625
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z21549
A:Accession: T34625
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-277 <SAU>
A:Cross-references: UNIPROT:Q9XAQ3; EMBL:AL078618; PIDN:CAB44532.1; GSPDB:GN00070; SCOEID:
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEID:SC10A7.22

Query Match      42.2%; Score 70; DB 2; Length 277;
Best Local Similarity 54.3%; Pred. No. 2;
Matches 19; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

QY 1 AKKYAKKEKAAYKKEAKAKAAEAAAKAEAYEA 35
   |||::|||::|||::|||::|||::|||
Db 93 AAKQAKTDLAEEKKAAEAAKKAEEAAKAEAAERA 127

RESULT 10
JC6552
DNA topoisomerase (EC 5.99.1.2) - slime mold (Physarum polycephalum)
N:Alternate names: emega-protein; micking-closing enzyme; type I DNA topoisomerase
C:Species: Physarum polycephalum
C>Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 09-Jul-2004
C:Accession: JC6552
R:Czerwinski, R.M.; Lipniacki, A.; Staron, K.
Gene 209, 39-44, 1998
```

```
A>Title: cDNA cloning of Physarum polycephalum DNA topoisomerase I and expression analysis
A:Reference number: JC6552; MUID:98245940; PMID:9583949
A:Accession: JC6552
A:Molecule type: mRNA
A:Residues: 1-1015 <CZE>
A:Cross-references: UNIPROT:Q94705; GB:AF023910; GB:U63217; NID:g2642492; PIDN:AAC14193.1
C:Genetics:
A:Gene: top1
C:Superfamily: eukaryotic type I DNA topoisomerase
C:Keywords: isomerase
F:974/Active site: Tyr #status predicted

Query Match      42.2%; Score 70; DB 2; Length 1015;
Best Local Similarity 42.9%; Pred. No. 5.9;
Matches 18; Conservative 6; Mismatches 8; Indels 10; Gaps 1;

QY 2 KKYAKKEKAAYKKEK-----AKAKAAEAAAKAEAY 33
   |||::|||::|||::|||::|||::|||::|||
Db 293 KKEVKEDTAKDVKVKEVKETPKTTPAKRKAESSSESDF 334

RESULT 11
T34081
hypothetical protein C02F12.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T34081
R:Miller, N.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid C02F12.
A:Reference number: Z21473
A:Accession: T34081
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1130 <MIL>
A:Cross-references: UNIPROT:Q11102; EMBL:U41545; PIDN:AAA83190.1; CESP:C02F12.7
C:Genetics:
A:Gene: CESP:C02F12.7
A:Introns: 11/2; 118/1; 156/2; 239/2; 308/2; 458/3; 591/1; 640/3; 756/2; 790/2; 815/2;

Query Match      42.2%; Score 70; DB 2; Length 1130;
Best Local Similarity 51.4%; Pred. No. 6.4;
Matches 18; Conservative 6; Mismatches 5; Indels 6; Gaps 1;

QY 6 KKEAKAKAYKKEAKAKA-----AEEAAKAEAYE 34
   |||::|||::|||::|||::|||::|||
Db 1016 KKEEASEKRTKAEKAEKAEKAEKAEKAEKAE 1050

RESULT 12
JV0057
tolA protein - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: JV0057; B64810
R:Levengood, S.K.; Webster, R.E.
J. Bacteriol. 171, 6600-6609, 1989
A>Title: Nucleotide sequences of the tolA and tolB genes and localization of their products in Escherichia coli
A:Reference number: JV0057; MUID:90078104; PMID:2687247
A:Accession: JV0057
A:Molecule type: DNA
A:Residues: 1-421 <LEV>
A:Cross-references: UNIPROT:P19934; GB:M28232; NID:g148018; PIDN:AAA24683.1; PID:g148019
A:Experimental source: strain JMW105
A>Note: the authors translated the initiation codon GTG for residue 1 as Val
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Coi
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B64810
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 39.2827 Seconds  
(without alignments)  
512.646 Million cell updates/sec

Title: US-10-792-311-1  
Perfect score: 166  
Sequence: 1 AKKYAKKEAKKAYKKEAKAKAAKAAKAAAYEA 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_02:.\*  
1: uniprot\_sprot:.\*  
2: uniprot\_trembl:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	50.6	384	1	TMPB TREPH
2	82	49.4	157	1	RS16 CORDI
3	82	49.4	157	2	CAE50058
4	81	48.8	372	2	Q9WXX1
5	81	48.8	372	2	Q88N16
6	80.5	48.5	461	2	Q68124
7	80	48.2	117	2	Q89DA0
8	79.5	47.9	98	2	Q8WQ44
9	79.5	47.9	111	2	Q8T9R3
10	78.5	47.3	105	2	Q6N503
11	78.5	47.3	105	2	CAE28621
12	78	47.0	395	2	Q6D7F3
13	78	47.0	441	2	Q6N9X8
14	78	47.0	441	2	CAE27215
15	77	46.4	140	1	RS16 SYNXP
16	77	46.4	275	1	MST3 DROHY
17	77	46.4	395	2	Q937K4
18	76	45.8	347	1	TOLA PSEAE
19	75	45.2	244	2	Q9AJX2
20	74.5	44.9	822	2	Q7S5A8
21	74.5	44.9	822	2	CAFO5892
22	74	44.6	189	2	Q8MYC2
23	74	44.6	372	1	TOLA HAEIN
24	73	44.0	156	2	Q8P6U0
25	72	43.4	140	2	Q6MNC2
26	72	43.4	140	2	CAE79230
27	72	43.4	153	1	RS16 BIFLO
28	72	43.4	210	1	H1 LYTFI
29	72	43.4	329	2	Q88RD8
30	71.5	43.1	104	2	Q89JH2
31	71.5	43.1	208	1	DBH_MYCSM

32	71.5	43.1	211	1	H1B_STRPU	P15869 strongyloce
33	71.5	43.1	297	2	Q7QC14	Q7GC14 anopheles g
34	71	42.8	168	1	RS16 COREF	Q8fp30 corynebacte
35	71	42.8	356	2	Q87Y39	Q87Y39 pseudomonas
36	71	42.8	373	2	Q7W477	Q7W477 bordetella
37	71	42.8	604	2	Q6F298	Q6F298 mesoplasma
38	71	42.8	625	2	Q87V67	Q87V67 pseudomonas
39	71	42.8	936	2	Q7UMV5	Q7UMV5 rhodospirell
40	70.5	42.5	44	2	Q711Y2	Q711Y2 lactobacill
41	70.5	42.5	44	2	AAQ06805	AAQ06805 lactobacil
42	70.5	42.5	445	2	Q9P3Q8	Q9P3Q8 neurospora
43	70	42.2	140	2	Q7YSZ7	Q7YSZ7 rhodnius pr
44	70	42.2	206	2	Q7M409	Q7M409 parechinus
45	70	42.2	277	2	Q9XAQ3	Q9XAQ3 streptomyce

ALIGNMENTS

RESULT 1

ID	TMPB	TREPH	STANDARD	PRT	384 AA.
AC	P29720				
DT	01-APR-1993	(Rel. 25, Created)			
DT	01-APR-1993	(Rel. 25, Last sequence update)			
DT	01-OCT-2004	(Rel. 45, Last annotation update)			
DE	Treponemal membrane protein B precursor (Antigen tmpB).				
GN	Name=tmpB;				
OS	Treponema phagedenis.				
OC	Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.				
OX	NCBI_TaxID=162;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Kazan 5;				
RX	MEDLINE=91372983; PubMed=1894368;				
RA	Yelton D.B., Limberger R.J., Curci K., Malinosky-Rummell F.,				
RA	Slivinsky L., Schouls L.M., van Embden J.D.A., Charon N.W.;				
RT	"Treponema phagedenis encodes and expresses homologs of the Treponema				
RT	pallidum tmpA and tmpB proteins.";				
RL	Infect. Immun. 59:3685-3693(1991).				
CC	-!- FUNCTION: Tmp may serve as a porin or transport protein for large				
CC	molecules.				
CC	-!- SUBCELLULAR LOCATION: Outer membrane-associated.				
CC	-!- SIMILARITY: To T.pallidum tmpB.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL	M58563	AAA27480.1	-	
DR	PIR	B43592	B43592		
DR	InterPro	IPR008941	TPR-like		
KW	Antigen; Outer membrane; Repeat; Signal.				
FT	SIGNAL	1	21	Potential.	
FT	CHAIN	22	384	Treponemal membrane protein B.	
FT	DOMAIN	151	235	17 X 5 AA tandem repeats of K-A-A-[AKR] -	
FT				[ED].	
FT	REPEAT	151	155	1-1.	
FT	REPEAT	156	160	1-2.	
FT	REPEAT	161	165	1-3.	
FT	REPEAT	166	170	1-4.	
FT	REPEAT	171	175	1-5.	
FT	REPEAT	176	180	1-6.	
FT	REPEAT	181	185	1-7.	
FT	REPEAT	186	190	1-8.	
FT	REPEAT	191	195	1-9.	
FT	REPEAT	196	200	1-10.	
FT	REPEAT	201	205	1-11.	
FT	REPEAT	206	210	1-12.	

FT	REPEAT	211	215	1-13.
FT	REPEAT	216	220	1-14.
FT	REPEAT	221	225	1-15.
FT	REPEAT	226	230	1-16.
FT	REPEAT	231	235	1-17.
FT	DOMAIN	236	288	6 X 8 AA tandem repeats of [EA]-A-A-R-X-A-A-E.
FT	REPEAT	236	243	2-1.
FT	REPEAT	245	252	2-2.
FT	REPEAT	254	261	2-3.
FT	REPEAT	263	270	2-4.
FT	REPEAT	272	279	2-5.
FT	REPEAT	281	288	2-6.
SQL	SEQUENCE	384 AA;	42677 MW;	6E94CBC74294DE8C CRC64;

Query Match 50.6%; Score 84; DB 1; Length 384;  
 Best Local Similarity 64.9%; Pred. No. 0.47;  
 Matches 24; Conservative 5; Mismatches 6; Indels 2; Gaps 2;

Qy	1	AKKYAKKEKAAK-KAYKKEAKAKAA-EAAAKEAAEAA 35
		:
Db	218	AREWAAKEKAADKAAKEEAARAAEAAAKAAEAA 254

RESULT 2

RS16	CORDI	STANDARD;	PRT;	157 AA.
ID	RS16_CORDI			
AC	P62228;			
DT	05-JUL-2004 (Rel. 44, Created)			
DT	05-JUL-2004 (Rel. 44, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	30S ribosomal protein S16.			
GN	Name=rpS; OrderedLocusNames=DIP1532;			
OS	Corynebacterium diphtheriae.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Corynebacterineae; Corynebacteriaceae; Corynebacterium.			
NCBI_TaxID=1717;				
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Biotype gravis / NCTC 13129;			
RX	MEDLINE=42965443; PubMed=14602910;			
RA	Cerdeno-Tarraga A.-M., Efstathiou A., Dover L.G., Holden M.T.G.,			
RA	Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,			
RA	De Zoysa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,			
RA	Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,			
RA	Rabinowitz E., Rutherford K.W., Thomson N.R., Unwin L.,			
RA	Whitehead S., Barrall B.G., Farkhill J.;			
RT	"The complete genome sequence and analysis of Corynebacterium			
RT	diphtheriae NCTC13129."			
RL	Nucleic Acids Res. 31:6516-6523(2003).			
CC	-1- SIMILARITY: Belongs to the S16P family of ribosomal proteins.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcements">http://www.isb-sib.ch/announcements</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
CC	EMBL; BX248358; CAB50058.1; -.			
DR	HAMAP; MF 00385; -1.			
DR	PROSITE; PS00732; RIBOSOMAL_S16; FALSE_NEG.			
KW	Complete proteome; Ribosomal protein.			
SQL	SEQUENCE 157 AA; 16985 MW; BB3FC04EEAD9708A CRC64;			

Query Match 49.4%; Score 82; DB 1; Length 157;  
 Best Local Similarity 64.5%; Pred. No. 0.34;  
 Matches 20; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy	2	KKYAKKEKAAK-KAYKKEAKAKAAEAAAKEAA 32
		:
Db	126	KKKAKKAAKAAEAAEAAAKAAEAPAEAA 156

[illegible]

```

RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RL Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005962; BAC52804.1; -.
KW Complete proteome.
SQ SEQUENCE 117 AA; 11800 MW; 1DC651CC17A1F2B5 CRC64;

Query Match 48.2%; Score 80; DB 2; Length 117;
Best Local Similarity 56.8%; Pred. No. 0.43;
Matches 21; Conservative 6; Mismatches 8; Indels 2; Gaps 1;

QY 1 AKKYAKK--EKAAYKAYKKEAKAKAAEAAAEAAEAA 35
   ||||| :||||: ||||| : ||||| : ||||| :
Db 21 AKKTAKKATKKAARKSAKSAKSAKSAKSAKSAKSAK 57

RESULT 8
Q8WQ44 PRELIMINARY; PRT; 98 AA.
AC Q8WQ44;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE 18 kDa nuclear protein.
GN Name=lnp18;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22326343; PubMed=12438377;
RA Papageorgiou F.T., Soteriadou K.P.;
RT "Expression of a novel Leishmania gene encoding a histone H1-like
RT protein in Leishmania major modulates parasite infectivity in vitro.";
RL Infect. Immun. 70:6976-6986(2002).
DR EMBL; AJ237814; CAD21431.1; -.
DR GO; GO:000786; C:nucleus; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006334; P:nucleosome assembly; IEA.
DR InterPro; IPR005819; Histone_H5.
DR PRINTS; PR00624; HISTONEH5.
KW Nuclear protein.
SQ SEQUENCE 98 AA; 9999 MW; 0A4AB93089D6C261 CRC64;

Query Match 47.9%; Score 79.5; DB 2; Length 98;
Best Local Similarity 67.6%; Pred. No. 0.41;
Matches 23; Conservative 3; Mismatches 5; Indels 3; Gaps 2;

QY 1 AKKYAKK--EKAAYKAYKKEAKAKAAEAAAEAAEAA 32
   ||||| :||||: ||||| : ||||| : ||||| :
Db 63 AKKVAKKPAKKAAPKAPKAPKAK-KAACKAAKAA 95

RESULT 9
Q8T9R3 PRELIMINARY; PRT; 111 AA.
AC Q8T9R3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE 18 kDa nuclear protein.
OS Leishmania infantum.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5671;
RN [1]
RP SEQUENCE FROM N.A.
RA Papageorgiou F., Soteriadou K.;
RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF469106; AAL76335.1; -.
DR GO; GO:000786; C:nucleus; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.

```

```

DR GO; GO:0006334; P:nucleosome assembly; IEA.
DR InterPro; IPR005819; Histone_H5.
DR PRINTS; PR00624; HISTONEH5.
KW Nuclear protein.
SQ SEQUENCE 111 AA; 11162 MW; 16168F3B54960E83 CRC64;

Query Match 47.9%; Score 79.5; DB 2; Length 111;
Best Local Similarity 67.6%; Pred. No. 0.46;
Matches 23; Conservative 3; Mismatches 5; Indels 3; Gaps 2;

QY 1 AKKYAKK--EKAAYKAYKKEAKAKAAEAAAEAAEAA 32
   ||||| :||||: ||||| : ||||| : ||||| :
Db 76 AKKVAKKPAKKAAPKAPKAPKAK-KAACKAAKAA 108

RESULT 10
Q6N503 PRELIMINARY; PRT; 105 AA.
AC Q6N503;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=RPA3180;
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R., Peres C.,
RA Harrison J.L., Hanson T.E., Bobst C., Torres Y.Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodopseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572603; CAE28621.1; -.
DR InterPro; IPR005819; Histone_H5.
DR PRINTS; PR00624; HISTONEH5.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 105 AA; 11042 MW; CEDB59B3D937B980 CRC64;

Query Match 47.3%; Score 78.5; DB 2; Length 105;
Best Local Similarity 63.2%; Pred. No. 0.56;
Matches 24; Conservative 4; Mismatches 7; Indels 3; Gaps 2;

QY 1 AKKYAKKKA-AKKAYK--KEAKAKAAEAAAEAAEAA 35
   ||||| :||||: ||||| : ||||| : ||||| :
Db 33 AKKVAKKPAKKAAPKAPKAPKAK-KAACKAAKAAKAA 70

RESULT 11
CAE28621 PRELIMINARY; PRT; 105 AA.
AC CAE28621;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN RPA3180.
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,

```

```

RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodospseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572603; CAE28621.1; -.
KW Hypothetical protein.
SQ SEQUENCE 105 AA; 11042 MW; CEDB59B3D937E980 CRC64;

Query Match 47.3%; Score 78.5; DB 2; Length 105;
Best Local Similarity 63.2%; Pred. No. 0.56;
Matches 24; Conservative 4; Mismatches 7; Indels 3; Gaps 2;

QY 1 AKKYAKKEKA-AKKAYK--KEAKAKAAEAEEAAKEAAVEA 35
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 33 AKKVAKVKAGAKAAKATKKGAKKAQAAKKAQAAKKA 70

RESULT 12
Q6D7F3 PRELIMINARY; PRT; 395 AA.
ID AC Q6D7F3;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Tola protein.
GN Name=tola; ORFNames=ECA1372;
OS Erwinia carotovora subsp. atroseptica SCRI1043.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=218491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCRI1043;
RA Bell K.S., Sebaihia M., Pritchard L., Holden M., Hyman L.J.,
RA Aleva M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Barrell B.G., Parkhill J., Toth I.K.,
RA Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX950851; CAG74282.1; -.
SQ SEQUENCE 395 AA; 41698 MW; 2DB1DAA64E984199 CRC64;

Query Match 47.0%; Score 78; DB 2; Length 395;
Best Local Similarity 57.1%; Pred. No. 2.1;
Matches 20; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 AKKYAKKEKA-AKKAYKKEAKAKAAEAEEAAKEAAVEA 35
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 173 AKQVAAAAAQQAEAEAKAAEAQKQAAETA 207

RESULT 13
Q6N8X8 PRELIMINARY; PRT; 441 AA.
ID AC Q6N8X8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE OmpA/MotB domain, possible porin precursor.
GN OrderedLocusNames=RPAL1774;
OS Rhodospseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodospseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA
```

```

RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodospseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
CC -!- SIMILARITY: Belongs to the ompA family.
DR EMBL; BX572598; CAE27215.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005741; C:mitochondrial outer membrane; IEA.
DR GO; GO:0015288; F:porin activity; IEA.
DR InterPro; IPR006664; Bac OmpA.
DR InterPro; IPR006665; OmpA/MotB.
DR Pfam; PF00691; OmpA; 1.
DR PRINTS; PRO1021; OMPADOMAIN.
DR ProDom; PD000930; OmpA/MotB; 1.
KW Complete proteome; Porin; Signal.
FT SIGNAL 1 27 Potential.
SQ SEQUENCE 441 AA; 44811 MW; F6BB66A4AFE183A0 CRC64;

Query Match 47.0%; Score 78; DB 2; Length 441;
Best Local Similarity 57.1%; Pred. No. 2.3;
Matches 20; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 AKKYAKKEKA-AKKAYKKEAKAKAAEAEEAAKEAAVEA 35
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 259 AKEVAAKEAAKAAEAEEAAKQAAADEAAKKADEA 293

RESULT 14
CAE27215 PRELIMINARY; PRT; 441 AA.
ID AC CAE27215;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE OmpA/MotB domain, possible porin precursor.
GN RPAL1774.
OS Rhodospseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodospseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodospseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572598; CAE27215.1; -.
KW Porin; Signal.
FT SIGNAL 1 27 Potential.
SQ SEQUENCE 441 AA; 44811 MW; F6BB66A4AFE183A0 CRC64;

Query Match 47.0%; Score 78; DB 2; Length 441;
Best Local Similarity 57.1%; Pred. No. 2.3;
Matches 20; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 AKKYAKKEKA-AKKAYKKEAKAKAAEAEEAAKEAAVEA 35
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 259 AKEVAAKEAAKAAEAEEAAKQAAADEAAKKADEA 293

RESULT 15
RS16 SYNXP STANDARD; PRT; 140 AA.
ID RS16 SYNXP
AC Q7TTU5;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE 30S ribosomal protein S16.
```

GN	Name: rpsP; Ordered locus Names = SYNW1617;
OC	Synechococcus sp. (strain WH8102).
OC	Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX	NCBI_TaxID=84588;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
RA	Palenik B., Brahamsha B., Larimer F.W., Land M.L., Hauser L.,
RA	Chapin P., Lamerdin J.E., Regala W., Allen E.E., McCarran J.,
RA	Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT	"The genome of a motile marine Synechococcus.";
RL	Nature 424:1037-1042(2003).
CC	-1- SIMILARITY: Belongs to the S16P family of ribosomal proteins.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see <a href="http://www.1ab-sib.ch/announcement/">http://www.1ab-sib.ch/announcement/</a>
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).
CC	-----
DR	EMBL; BX569693; CAE08132.1; -.
DR	HAWK; MF_00385; -, 1
DR	InterPro; IPR000307; Ribosomal_S16.
DR	Pfam; PF00886; Ribosomal_S16; 1.
DR	ProDom; PD003791; Ribosomal_S16; 1.
DR	TIGRFAMS; TIGR00002; S16; 1.
DR	PROSITE; PS00732; RIBOSOMAL_S16; 1.
KW	Complete proteome; Ribosomal protein.
SQ	SEQUENCE 140 AA; 15194 MW; 99C48AE8D709EA42 CRC64;
Query Match	46.4%; Score 77; DB 1; Length 140;
Best Local Similarity	54.3%; Pred. No. 1;
Matches	19; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
Qy	1 AKYAKKEKAAYKAYKKEAKAKAAEAANKAAAYEA 35
Db	90 AKQAATKEEAKQAQKAAEAKAAEAFAAFAAFAA 124

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 38.692 Seconds  
(without alignments)  
324.499 Million cell updates/sec

Title: US-10-792-311-1

Perfect score: 166

Sequence: 1 AKKYAKKEKAAYKAYKEAKAKAAAEAAKEAAEYA 35

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_238sep04.\*  
1: geneseqp1980s.\*  
2: geneseqp1990s.\*  
3: geneseqp2000s.\*  
4: geneseqp2001s.\*  
5: geneseqp2002s.\*  
6: geneseqp2003as.\*  
7: geneseqp2003bs.\*  
8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	166	100.0	35	3 AAY82571	Aay82571 Copolymer
2	108	65.1	45	3 AAY82572	Aay82572 Copolymer
3	108	65.1	86	3 AAY82576	Aay82576 Copolymer
4	105.5	63.6	66	3 AAY82574	Aay82574 Copolymer
5	105	63.3	109	3 AAY82577	Aay82577 Copolymer
6	104.5	63.0	56	3 AAY82573	Aay82573 Copolymer
7	104.5	63.0	77	3 AAY82575	Aay82575 Copolymer
8	86.5	52.1	427	8 AAY82575	Aag70868 C albican
9	86.5	52.1	427	8 ADP99051	Adp99051 C. albica
10	82	49.4	157	6 ABU26064	Abu26064 Protein e
11	81	48.8	372	6 ABU40185	Abu40185 Protein e
12	76	45.8	154	2 AAR06445	Aar06445 Recombina
13	76	45.8	347	6 ABU38313	Abu38313 Protein e
14	76	45.8	347	6 ABJ18771	Abj18771 Pseudomon
15	76	45.8	407	7 ABO80835	Abc80835 Pseudomon
16	74.5	44.9	80	5 ABG71044	Abg71044 Tumour ne
17	74	44.6	372	5 AEG80418	Abg80418 Haemophil
18	74	44.6	372	7 ABO23507	Abc23507 Haemophil
19	73.5	44.3	40	3 AAB08170	Aab08170 Peptide m
20	72	43.4	80	7 ADI10657	Adi10657 Structura
21	72	43.4	80	7 ADI10656	Adi10656 Structura
22	72	43.4	80	7 ADI10606	Adi10606 Structura
23	72	43.4	80	7 ADI10607	Adi10607 Structura
24	72	43.4	80	8 ADKI5676	Adki5676 Library f
25	72	43.4	80	8 ADKI5675	Adki5675 Library f

26	72	43.4	80	8 ADKI5626	Adki5626 Nucleatin
27	72	43.4	80	8 ADKI5625	Adki5625 Nucleatin
28	72	43.4	85	7 ADE10608	Adel10608 Structura
29	72	43.4	85	7 ADE10658	Adel10658 Structura
30	72	43.4	85	8 ADKI5627	Adki5627 Nucleatin
31	72	43.4	85	8 ADKI5677	Adki5677 Library f
32	72	43.4	153	5 ABP65446	Abp65446 Bifidobac
33	72	43.4	329	6 ABU39775	Abu39775 Protein e
34	71	42.8	31	8 ADN11701	Adn11701 Peptide l
35	71	42.8	106	2 AAR06446	Aar06446 Recombina
36	71	42.8	336	6 ABU42038	Abu42038 Protein e
37	69.5	41.9	145	4 AAM25508	Aam25508 Human pro
38	69.5	41.9	165	4 AAG91997	Aag91997 C glutami
39	69.5	41.9	421	6 ABU28559	Abu28559 Protein e
40	69	41.6	741	4 ABB62028	Abb62028 Drosophi
41	67.5	40.7	32	2 AAR90180	Aar90180 Polycatio
42	67.5	40.7	32	2 AAW06686	Aaw06686 Protamine
43	67.5	40.7	32	3 AAB08168	Aab08168 Peptide m
44	67.5	40.7	33	2 AAR90181	Aar90181 Polycatio
45	67.5	40.7	33	2 AAW06688	Aaw06688 Protamine

ALIGNMENTS

RESULT 1

AAY82571

ID AAY82571 standard; peptide; 35 AA.

AC AAY82571;

DT 28-JUL-2000 (first entry)

DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:1.

Copolymer; molecular weight marker; TV-marker; immune disease;  
glutiramer acetate; autoimmune disease; antiarthritic; neuroprotective;  
osteopathic; immunosuppressive; antithyroid; antiinflammatory;  
antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;  
antianemic; immunosuppressive; demyelinating disease; osteoarthritis;  
inflammatory condition; multiple sclerosis; rheumatoid arthritis;  
Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;  
diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;  
Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;  
pemphigus vulgaris; systemic lupus erythematosus.

OS Unidentified.

XX WO200018794-A1.

XX 06-APR-2000.

XX 24-SEP-1999; 99WO-US022402.

XX 25-SEP-1998; 98US-0101693P.

XX (YEDA ) YEDA RES & DEV CO LTD.

XX (TEVA-) TEVA PHARM USA INC.

XX Gad A, Lis D;

XX WPI; 2000-317499/27.

XX Copolymer 1 related polypeptides used as molecular weight markers for  
glutiramer acetate and for treatment and prevention of immune diseases.  
Claim 10; Page 14; 72pp; English.

XX AAY82571 to AAY82577 represent specifically claimed copolymer molecular  
weight TV-marker polypeptides from the present invention. The present  
invention describes polypeptides (I) for determining the molecular weight  
of a copolymer (CP), which has an identified molecular weight and an  
amino acid composition corresponding to the copolymer. The polypeptides

CC of the invention are used as molecular weight markers for glatiramer  
 CC acetate related tetrapolymers. The polypeptides may also be used for  
 CC treating and preventing immune diseases in a mammal. Autoimmune diseases  
 CC which may be treated include either cell-mediated or antibody-mediated  
 CC diseases. Such diseases include arthritic conditions, demyelinating  
 CC arthritis, osteoarthritis, autoimmune haemolytic anaemia, rheumatoid  
 CC oophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's  
 CC disease, chronic immune thrombocytopaenia purpura, colitis, contact  
 CC sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's  
 CC syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,  
 CC psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-  
 CC mediated diseases which can be treated include host-versus-graft disease,  
 CC graft-versus-host disease, and delayed-type hypersensitivity. The  
 CC polypeptides of the invention have defined molecular weights and physical  
 CC properties which are analogous to glatiramer acetate molecules, which  
 CC makes them ideal for use as molecular weight markers  
 XX  
 SQ Sequence 35 AA;

Query Match 100.0%; Score 166; DB 3; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 2e-12; Mismatches 0; Indels 0; Gaps 0;  
 Matches 35; Conservative 0;

QY 1 AKKYAKKEKAAKAYKKEAKAKAAEAKEAAEAYEA 35  
 |||||  
 DB 1 AKKYAKKEKAAKAYKKEAKAKAAEAKEAAEAYEA 35

RESULT 2  
 AAY82572  
 ID AAY82572 standard; peptide; 45 AA.

AC AAY82572;  
 XX  
 XX  
 DT 28-JUL-2000 (first entry)  
 DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:2.  
 XX  
 XX Copolymer; molecular weight marker; TV-marker; immune disease;  
 KW glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;  
 KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;  
 KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;  
 KW antianaemic; immunosuppressive; demyelinating disease; osteoarthritis;  
 KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;  
 KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;  
 KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;  
 KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;  
 KW pemphigus vulgaris; systemic lupus erythematosus.

XX Unidentified.  
 XX WO200018794-A1.

XX 06-APR-2000.

XX 24-SEP-1999; 99WO-US022402.

XX 25-SEP-1998; 98US-0101693P.

XX (YEDA ) YEDA RES & DEV CO LTD.  
 XX (TEVA-) TEVA PHARM USA INC.

XX Gad A, Lis D;

XX WPI; 2000-317499/27.

XX Copolymer 1 related polypeptides used as molecular weight markers for  
 PT glatiramer acetate and for treatment and prevention of immune diseases.

PS Claim 10; Page 14; 72pp; English.

XX AAY82571 to AAY82577 represent specifically claimed copolymer molecular

CC weight TV-marker polypeptides from the present invention. The present  
 CC invention describes polypeptides (I) for determining the molecular weight  
 CC of a copolymer (CP), which has an identified molecular weight and an  
 CC amino acid composition corresponding to the copolymer. The polypeptides  
 CC of the invention are used as molecular weight markers for glatiramer  
 CC acetate related tetrapolymers. The polypeptides may also be used for  
 CC treating and preventing immune diseases in a mammal. Autoimmune diseases  
 CC which may be treated include either cell-mediated or antibody-mediated  
 CC diseases. Such diseases include arthritic conditions, demyelinating  
 CC diseases and inflammatory conditions, e.g. multiple sclerosis, rheumatoid  
 CC arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune  
 CC oophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's  
 CC disease, chronic immune thrombocytopaenia purpura, colitis, contact  
 CC sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's  
 CC syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,  
 CC psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-  
 CC mediated diseases which can be treated include host-versus-graft disease,  
 CC graft-versus-host disease, and delayed-type hypersensitivity. The  
 CC polypeptides of the invention have defined molecular weights and physical  
 CC properties which are analogous to glatiramer acetate molecules, which  
 CC makes them ideal for use as molecular weight markers  
 XX  
 SQ Sequence 45 AA;

Query Match 65.1%; Score 108; DB 3; Length 45;  
 Best Local Similarity 64.4%; Pred. No. 1.4e-05;  
 Matches 29; Conservative 1; Mismatches 5; Indels 10; Gaps 2;

QY 1 AKKYAKKEKA--AKKAYK-----KEAKAKAAEAKEAAEAYEA 35  
 |||||  
 DB 1 AKKYAKKAKAKAKAYKAAEAKEAKAAYKEKAAEAKEAAEAYEA 45

RESULT 3  
 AAY82576  
 ID AAY82576 standard; peptide; 86 AA.

AC AAY82576;

DT 28-JUL-2000 (first entry)

DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:6.

XX Copolymer; molecular weight marker; TV-marker; immune disease;  
 KW glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;  
 KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;  
 KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;  
 KW antianaemic; immunosuppressive; demyelinating disease; osteoarthritis;  
 KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;  
 KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;  
 KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;  
 KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;  
 KW pemphigus vulgaris; systemic lupus erythematosus.

XX Unidentified.

XX WO200018794-A1.

XX 06-APR-2000.

XX 24-SEP-1999; 99WO-US022402.

XX 25-SEP-1998; 98US-0101693P.

XX (YEDA ) YEDA RES & DEV CO LTD.  
 XX (TEVA-) TEVA PHARM USA INC.

XX Gad A, Lis D;

XX WPI; 2000-317499/27.

XX Copolymer 1 related polypeptides used as molecular weight markers for  
 PT glatiramer acetate and for treatment and prevention of immune diseases.



XX Claim 10; Page 14; 72pp; English.

XX AAY82571 to AAY82577 represent specifically claimed copolymer molecular

XX weight TV-marker polypeptides from the present invention. The present

XX invention describes polypeptides (I) for determining the molecular weight

XX of a copolymer (CP), which has an identified molecular weight and an

XX amino acid composition corresponding to the copolymer. The polypeptides

XX of the invention are used as molecular weight markers for glatiramer

XX acetate related tetrapolymers. The polypeptides may also be used for

XX treating and preventing immune diseases in a mammal. Autoimmune diseases

XX which may be treated include either cell-mediated or antibody-mediated

XX diseases. Such diseases include arthritic conditions, demyelinating

XX diseases and inflammatory conditions, e.g. multiple sclerosis, rheumatoid

XX arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune

XX oophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's

XX disease, chronic immune thrombocytopaenia purpura, colitis, contact

XX sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's

XX syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,

XX mediated diseases which can be treated include host-versus-graft disease,

XX graft-versus-host disease, and delayed-type hypersensitivity. The

XX polypeptides of the invention have defined molecular weights and physical

XX properties which are analogous to glatiramer acetate molecules, which

XX makes them ideal for use as molecular weight markers

XX Sequence 86 AA;

SQ Query Match 65.1%; Score 108; DB 3; Length 86;

Best Local Similarity 57.4%; Pred. No. 2.8e-05;

Matches 27; Conservative 2; Mismatches 6; Indels 12; Gaps 1;

Qy 1 AKKYAKKEKAAYKAY-----KKEAKAKAAEAAYEA 35

Db 40 AKKYAKAAKAEKYEAAAEAKYKAEAAKAYKAEAAKAAKAEAAEA 86

RESULT 4

AAY82574

ID AAY82574 standard; peptide; 66 AA.

XX AAY82574;

AC AAY82574;

XX 28-JUL-2000 (first entry)

DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:4.

XX Copolymer; molecular weight marker; TV-marker; immune disease;

KW glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;

KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;

KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;

KW antianaemic; immunosuppressive; demyelinating disease; osteoarthritis;

KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;

KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;

KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;

KW pemphigus vulgaris; systemic lupus erythematosus.

XX Unidentified.

XX WO200018794-A1.

FN WO200018794-A1.

PD 06-APR-2000.

XX 24-SEP-1999; 99WO-US022402.

PR 25-SEP-1998; 98US-0101693P.

XX (YEDA ) YEDA RES & DEV CO LTD.

PA (TEVA-) TEVA PHARM USA INC.

XX Gad A, Lis D;

XX

DR WPI; 2000-317499/27.

XX Copolymer 1 related polypeptides used as molecular weight markers for

XX glatiramer acetate and for treatment and prevention of immune diseases.

PS Claim 10; Page 14; 72pp; English.

XX AAY82571 to AAY82577 represent specifically claimed copolymer molecular

XX weight TV-marker polypeptides from the present invention. The present

XX invention describes polypeptides (I) for determining the molecular weight

XX of a copolymer (CP), which has an identified molecular weight and an

XX amino acid composition corresponding to the copolymer. The polypeptides

XX of the invention are used as molecular weight markers for glatiramer

XX acetate related tetrapolymers. The polypeptides may also be used for

XX treating and preventing immune diseases in a mammal. Autoimmune diseases

XX which may be treated include either cell-mediated or antibody-mediated

XX diseases. Such diseases include arthritic conditions, demyelinating

XX diseases and inflammatory conditions, e.g. multiple sclerosis, rheumatoid

XX arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune

XX oophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's

XX disease, chronic immune thrombocytopaenia purpura, colitis, contact

XX sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's

XX syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,

XX psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-

XX mediated diseases which can be treated include host-versus-graft disease,

XX graft-versus-host disease, and delayed-type hypersensitivity. The

XX polypeptides of the invention have defined molecular weights and physical

XX properties which are analogous to glatiramer acetate molecules, which

XX makes them ideal for use as molecular weight markers

XX Sequence 66 AA;

SQ Query Match 63.6%; Score 105.5; DB 3; Length 66;

Best Local Similarity 43.9%; Pred. No. 4.2e-05;

Matches 29; Conservative 2; Mismatches 4; Indels 31; Gaps 1;

Qy 1 AKKYAKKEKAAYKAYKKEA-----KAKAAEA 29

Db 1 AKKYAKKEKAYAKAKAEAKAAKAAKAAKAAKAAKAAKAAKAAKAA 60

Qy 30 EAYEA 35

Db 61 EAYEA 66

RESULT 5

AAY82577

ID AAY82577 standard; peptide; 109 AA.

XX AAY82577;

AC AAY82577;

DT 28-JUL-2000 (first entry)

DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:7.

XX Copolymer; molecular weight marker; TV-marker; immune disease;

KW glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;

KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;

KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;

KW antianaemic; immunosuppressive; demyelinating disease; osteoarthritis;

KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;

KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;

KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;

KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;

KW pemphigus vulgaris; systemic lupus erythematosus.

XX Unidentified.

OS WO200018794-A1.

XX WO200018794-A1.

PN 06-APR-2000.

PD 24-SEP-1999; 99WO-US022402.

XX





Constructing a strain of diploid fungal cells in which both alleles of a gene are modified comprises modifying the alleles of a gene in the fungal cells by recombination using a gene disruption cassette and a promoter replacement fragment.

Claim 44; SEQ ID NO 7226; 163pp; English.

The invention relates to a novel method for constructing a strain of diploid fungal cells in which both alleles of a gene are modified. The method comprises modifying the alleles of a gene in diploid fungal cells by recombination using a gene disruption cassette and a promoter replacement fragment. The invention further comprises: assembling a collection of diploid fungal cells each of which comprises modified alleles of a different gene; a strain of diploid fungal cells comprising modified alleles of a gene, where the first allele of the gene is inactivated by a gene disruption cassette comprising a nucleotide sequence encoding an expressible selectable marker; and the expression of the second allele of the gene is regulated by a heterologous promoter that is operably linked to the coding region of the second allele of the gene, and where the gene encodes the polypeptide mentioned above; a collection of diploid fungal strains comprising the diploid strains cited above, where substantially all the different genes that encode the above amino acid sequences are modified and are present in different diploid strains in the collection; a nucleic acid molecule microarray comprising nucleic acid molecules, where each nucleic acid molecule comprises a nucleotide sequence that is hybridizable to a target nucleotide sequence comprising any of the 310 nucleotide sequences listed in the specification (ADP98516-ADP98825); identifying a gene that is essential to the survival or growth of a fungus, that contributes to the virulence and/or pathogenicity of a fungus, or that contributes to the resistance of a diploid fungus to an antifungal agent; identifying an antifungal agent that inhibits the growth of a diploid fungus, or a therapeutic agent for treatment of a mammalian disease; correlating changes in the levels of proteins or gene transcripts with the inhibition of growth or proliferation of a diploid fungal cell; a purified or isolated nucleic acid molecule comprising a nucleotide sequence encoding a gene product required for proliferation of Candida albicans, where the gene product consists of any of the above-mentioned amino acid sequences; a vector comprising a promoter operably linked to the nucleic acid molecule cited above; a host cell containing the vector; a purified or isolated polypeptide comprising any of the 61 amino acid sequences given in the specification (ADP96718-ADP96778); a fusion protein comprising a fragment of a first polypeptide fused to a second polypeptide, the fragment consisting of at least 6 consecutive residues of any of ADP98826-ADP99135; or producing a polypeptide; identifying a compound which modulates the activity of a gene product encoded by a nucleic acid comprising any of ADP98516-ADP98825; eliciting an immune response in an animal; a strain of Candida albicans, where a first allele of a gene comprising any of ADP98516-ADP98825 is inactive and a second allele of the gene is under the control of a heterologous promoter; identifying a compound or binding partner that binds to the polypeptide comprising any of ADP98826-ADP99135, or its fragment; identifying a compound having the ability to inhibit growth or proliferation of Candida albicans; inhibiting growth or proliferation of Candida albicans cells; manufacturing an antimycotic compound; treating an infection of a subject by Candida albicans; preventing or containing contamination of an object by Candida albicans, or for preventing or inhibiting formation on a surface of a biofilm comprising Candida albicans; a pharmaceutical composition comprising a therapeutic amount of an agent which reduces the activity or level of a gene product encoded by a nucleic acid comprising any of ADP98516-ADP98825 in a pharmaceutical carrier; an antibody preparation which binds the polypeptide; methods for evaluating a compound against a target gene product encoded by any of ADP98516-ADP98825; identifying an antimycotic compound; a computer or a computer readable medium that comprises at least one of the nucleotide sequences mentioned in the specification or at least one amino acid sequence selected from ADP98826-ADP99135; a method assisted by a computer for identifying a putatively essential gene of a fungus; and a protein array comprising proteins, where at least one protein comprises an amino acid sequence or a portion of an amino acid sequence selected from ADP98516-ADP98825. The novel methods and compositions have fungicidal activity. The compositions may be used in gene therapy. The composition and methods are useful for drug screening purposes or for diagnosing, preventing or treating infections associated

CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 157 AA;

Query Match 49.4%; Score 82; DB 6; Length 157;  
 Best Local Similarity 64.5%; Pred. No. 0.054;  
 Matches 20; Conservative 3; Mismatches 8; Indels 0; Gaps 0;  
 QY 2 KKYAKKEAAKAYKKEAKAKAAEAAAEAA 32  
 |||||:||||:|||||:|||||:|||||  
 Db 126 KKKAKEAAKAAEAAEAAEAAEAAEAAEAA 156

RESULT 11  
 ABU040185  
 ID ABU040185 standard; protein; 372 AA.  
 XX  
 AC ABU040185;  
 DT 19-JUN-2003 (first entry)  
 XX  
 DE Protein encoded by Prokaryotic essential gene #25712.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX Pseudomonas putida.  
 OS WO200277183-A2.  
 PN  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002WO-US009107.

XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyekind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 DR WPI; 2003-029926/02.  
 DR N-PSDB; ACA44055.

XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 PS Claim 25; SEQ ID NO 68109; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 372 AA;

Query Match 48.8%; Score 81; DB 6; Length 372;  
 Best Local Similarity 56.1%; Pred. NO. 0.17;  
 Matches 23; Conservative 3; Mismatches 9; Indels 6; Gaps 1;  
 QY 1 AKKYA-----KKKKAACKKAYKKEAKAKAAEAAAEAAEAA 35  
 |||||:|||||:|||||:|||||:|||||:|||||  
 Db 160 AKKXADEAKKXAEAAEAAEAAEAAEAAEAAEAAEAAEAA 200

RESULT 12  
 AAR06445  
 ID AAR06445 standard; protein; 154 AA.  
 XX  
 AC AAR06445;  
 DT 25-MAR-2003 (revised)  
 DT 03-JAN-1991 (first entry)  
 XX  
 DE Recombinant copolymer 1-77, myelin basic protein analogue.  
 XX  
 KW Recombinant copolymer 1; COP-1-77; myelin basic protein; MBP;  
 KW immunological activity; autoimmune encephalomyelitis; multiple sclerosis.  
 XX  
 OS Synthetic.  
 XX  
 PN EP383620-A.  
 XX  
 PD 22-AUG-1990.  
 XX  
 PF 16-FEB-1990; 90EP-00301700.  
 XX  
 PR 17-FEB-1989; 89US-00312541.  
 PR 07-FEB-1990; 90US-00473845.  
 XX  
 XX (REPK ) REPLIGEN CORP.  
 PA  
 XX  
 PI Cook KS;  
 XX  
 DR WPI; 1990-255848/34.  
 DR N-PSDB; AAQ05664.  
 XX  
 PT Producing genes encoding random polymers of aminoacid(s) - for producing

PT recombinant polypeptide(s) with biological and/or immunological activity.  
 PS Disclosure; Fig 11; 25pp; English.

XX To improve the expression of rCOP-1 polypeptides in *E. coli*, genes coding  
 CC for rCOP-1-77 were subcloned from pREV 2.1 to pBG3-2deltaN (deposit: 20-  
 CC NOV-1984 US4691009, NRRL B-15910), a plasmid used to express Protein A.  
 CC The resulting plasmids encode fusion proteins consisting of beta-  
 CC glucuronidase, Protein A, and rCOP-1 sequences. A methionine residue  
 CC occurs between the Protein A and rCOP-1 sequences, originating from the  
 CC 5' linker sequence, in order that the COP-1 polypeptide may be cleaved  
 CC from the fusion protein. rCOP-1-77 contains oligonucleotide duplexes  
 CC in coding the following segments: YKK, EAE, KAK, AAK, and AAA. The N-  
 CC terminal alanine residue is left behind following CNBR cleavage of the  
 CC fusion protein. The product prevents or arrests experimental autoimmune  
 CC encephalomyelitis. They are used to prevent, arrest or control a  
 CC demyelinating disorder, e.g. multiple sclerosis. They may also be used as  
 CC additives to hair care products to confer beneficial effects on damaged  
 CC hair or as supplements for diets deficient in certain amino acids. See  
 CC also AAQ05665. (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 XX Sequence 154 AA;

Query Match 45.8%; Score 76; DB 2; Length 154;  
 Best Local Similarity 54.1%; Pred. No. 0.26;  
 Matches 20; Conservative 4; Mismatches 9; Indels 4; Gaps 1;

QY 2 KKYAKKEKAAKA---YKKAKEAKAAEAKEAYE 34  
 DB 24 KKKAKEAKKAKYKYYKKAEAKAAKAAKAAAYK 60

RESULT 13

ABU38313  
 ID ABU38313 standard; protein; 347 AA.

XX AC ABU38313;

XX 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #23840.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Pseudomonas aeruginosa.

XX WO20027183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR N-PSDB; AC442183.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 66237; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation; or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 347 AA;

Query Match 45.8%; Score 76; DB 6; Length 347;  
 Best Local Similarity 56.1%; Pred. No. 0.6;  
 Matches 23; Conservative 5; Mismatches 7; Indels 6; Gaps 2;

QY 1 AKKYAK---KEKAAKKAYKK---EAKAKAAEAKEAYEA 35  
 DB 155 AKKRAEDENKKUAEADAKKAAEDAKKAAEAKKKAAEA 195

RESULT 14

ABU18771

ID ABU18771 standard; protein; 347 AA.

XX AC ABU18771;

XX 27-FEB-2003 (first entry)

XX Pseudomonas aeruginosa biofilm formation-related protein #35.

XX Biofilm formation modulation; biofilm-associated disease;  
 KW cystic fibrosis; AIDS; middle ear infection; acne; periodontal disease;  
 KW catheter-associated infection; medical device-associated infection.

OS Pseudomonas aeruginosa.

XX WO200285295-A2.

XX 31-OCT-2002.

XX 19-APR-2002; 2002WO-US012532.

XX 20-APR-2001; 2001US-0285190P.

PR 24-OCT-2001; 2001US-0344142P.

XX (IOWA ) UNIV IOWA RES FOUND.

PA (HARD ) HARVARD COLLEGE.

XX Whiteley M, Banger MG, Lory S, Greenberg EP;

XX WPI; 2003-075601/07.

DR N-PSDB; ABT14593.

XX Identifying compound capable of modulating biofilm formation by

PT bacteria/bacterial antibiotic resistance, useful for treatment of biofilm

PT associated disease.

XX Claim 1; Page 119-120; 154pp; English.

XX The invention comprises a method for identifying a compound capable of

CC modulating biofilm formation by bacteria. The method of the invention is

CC useful for identifying a compound capable of modulating biofilm formation

CC by bacteria or modulating bacterial antibiotic resistance. The method of

CC the invention is also useful for diagnosing and treating a subject

CC (especially an immunocompromised human) that is afflicted with a biofilm-

CC associated disease or disorder, such as: cystic fibrosis; AIDS; middle

CC ear infections; acne; periodontal disease; catheter-associated infections

CC and medical device-associated infections. The present amino acid

CC sequence represents a protein that is used in the invention

XX

SQ Sequence 347 AA;

Query Match 45.8%; Score 76; DB 6; Length 347;

Best Local Similarity 56.1%; Pred. No. 0.6;

Matches 23; Conservative 5; Mismatches 7; Indels 6; Gaps 2;

QY 1 AKKYAK---KEKAAYK---EAKAKAAEAKEAAVEA 35

Db 155 AKKRAEAKKAAEDAKKAAEDAKKAAEDAKKAAEDAKKAAEA 195

RESULT 15

ABO80835

ID ABO80835 standard; protein; 407 AA.

XX ABO80835;

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #13010.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI: 2003-615309/58.

XX N-PSDB; ABD14406.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,

PT useful as molecular targets for diagnostics, prophylaxis and treatment of

PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 29581; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the

CC polynucleotides encoding them. The sequences are useful in diagnosis and

CC therapy of pathological conditions, as molecular targets for diagnostics,

CC prophylaxis and treatment of pathological conditions resulting from a

CC bacterial infection, for evaluating a compound, such as a polypeptide,

CC for the ability to bind a P. aeruginosa nucleic acid, as components of

CC effective antibacterial targets, as targets for antibacterial drugs,

CC

CC including anti-P. aeruginosa drugs, as templates for recombinant

CC production of P. aeruginosa-derived peptides or polypeptides, as target

CC components for diagnosis and/or treatment of P. aeruginosa-caused

CC infection, and in detection of P. aeruginosa sequences or other sequences

CC of Pseudomonas species using biochip technology. Sequences ABO67826-

CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The

CC sequence data for this patent did not form part of the printed

CC specification but was obtained in electronic format from USPTO at

XX seqdata.uspto.gov/sequence.html

SQ Sequence 407 AA;

Query Match 45.8%; Score 76; DB 7; Length 407;

Best Local Similarity 56.1%; Pred. No. 0.7;

Matches 23; Conservative 5; Mismatches 7; Indels 6; Gaps 2;

QY 1 AKKYAK---KEKAAYK---EAKAKAAEAKEAAVEA 35

Db 215 AKKRAEAKKAAEDAKKAAEDAKKAAEDAKKAAEDAKKAAEA 255

Search completed: December 14, 2004, 06:01:08

Job time : 46.692 secs

**This Page Blank (usp10)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:52:22 ; Search time 86.8354 Seconds  
(without alignments)  
143.965 Million cell updates/sec

Title: US-10-792-311-1

Perfect score: 166

Sequence: 1 AKKYAKKEKAAYKKEAKAAEAAAEAAAYEA 35

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	166	100.0	35	9	US-09-816-989A-1	Sequence 1, Appli
2	108	65.1	45	9	US-09-816-989A-2	Sequence 2, Appli
3	108	65.1	86	9	US-09-816-989A-6	Sequence 6, Appli
4	105.5	63.6	66	9	US-09-816-989A-4	Sequence 4, Appli
5	105	63.3	109	9	US-09-816-989A-7	Sequence 7, Appli
6	104.5	63.0	56	9	US-09-816-989A-3	Sequence 3, Appli
7	104.5	63.0	77	9	US-09-816-989A-5	Sequence 5, Appli
8	82	49.4	157	15	US-10-282-122A-53988	Sequence 53988, A
9	81	48.8	372	15	US-10-282-122A-68109	Sequence 68109, A
10	76	45.8	347	14	US-10-127-032-120	Sequence 120, App
11	76	45.8	347	15	US-10-282-122A-66237	Sequence 66237, A
12	74	44.6	372	10	US-09-820-843A-8	Sequence 8, Appli
13	74	44.6	372	16	US-10-467-421-16	Sequence 16, Appli

14	72	43.4	80	14	US-10-177-725-13	Sequence 13, Appli
15	72	43.4	80	14	US-10-177-725-14	Sequence 14, Appli
16	72	43.4	80	14	US-10-177-725-63	Sequence 63, Appli
17	72	43.4	80	14	US-10-177-725-64	Sequence 64, Appli
18	72	43.4	80	14	US-10-393-449-13	Sequence 13, Appli
19	72	43.4	80	14	US-10-393-449-14	Sequence 14, Appli
20	72	43.4	80	14	US-10-393-449-63	Sequence 63, Appli
21	72	43.4	80	14	US-10-393-449-64	Sequence 64, Appli
22	72	43.4	80	14	US-10-177-725-15	Sequence 15, Appli
23	72	43.4	85	14	US-10-177-725-65	Sequence 65, Appli
24	72	43.4	85	14	US-10-393-449-15	Sequence 15, Appli
25	72	43.4	85	14	US-10-393-449-65	Sequence 65, Appli
26	72	43.4	329	15	US-10-282-122A-67699	Sequence 67699, A
27	71	42.5	336	15	US-10-282-122A-69962	Sequence 69962, A
28	70.5	42.8	189	16	US-10-767-701-60774	Sequence 60774, A
29	70	42.2	1130	14	US-10-369-493-6751	Sequence 6751, A
30	69.5	41.9	146	15	US-10-296-115-1023	Sequence 1023, Ap
31	69.5	41.9	165	9	US-09-738-626-5751	Sequence 5751, Ap
32	69.5	41.9	421	15	US-10-282-122A-56483	Sequence 56483, A
33	67	40.4	258	14	US-10-156-761-9957	Sequence 9957, Ap
34	66.5	40.1	79	14	US-10-177-725-20	Sequence 20, Appli
35	66.5	40.1	79	14	US-10-393-449-20	Sequence 20, Appli
36	66.5	40.1	582	9	US-09-919-497-100	Sequence 100, App
37	66	39.8	120	16	US-10-767-701-45061	Sequence 45061, A
38	66	39.8	269	17	US-10-425-115-27374	Sequence 27374, A
39	66	39.8	526	15	US-10-282-122A-53742	Sequence 53742, A
40	66	39.8	685	14	US-10-369-493-3684	Sequence 3684, Ap
41	66	39.8	1593	15	US-10-282-122A-65262	Sequence 65262, A
42	65.5	39.5	79	14	US-10-177-725-16	Sequence 16, Appli
43	65.5	39.5	79	14	US-10-393-449-16	Sequence 16, Appli
44	65.5	39.5	212	15	US-10-282-122A-61735	Sequence 61735, A
45	65.5	39.5	214	14	US-10-229-567-27	Sequence 27, Appli

ALIGNMENTS

RESULT 1

US-09-816-989A-1  
; Sequence 1, Application US/09816989A  
; Patent No. US20020115103A1  
; GENERAL INFORMATION:  
; APPLICANT: Gad, Alexander  
; APPLICANT: Lis, Doris  
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARKI  
; FILE REFERENCE: 2609/60807-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/816,989A  
; PRIOR FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: 60/101,693  
; PRIOR FILING DATE: 1998-09-25  
; PRIOR APPLICATION NUMBER: PCT/US99/22402  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 1  
; LENGTH: 35  
; TYPE: PPT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide  
US-09-816-989A-1

Query Match 100.0%; Score 166; DB 9; Length 35;  
Best Local Similarity 100.0%; Pred. No. 1.8e-11;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKKYAKKEKAAYKKEAKAAEAAAEAAAYEA 35

Db 1 AKKYAKKEKAAYKKEAKAAEAAAEAAAYEA 35

RESULT 2

```
US-09-816-989A-2
; Sequence 2, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-2
Query Match 65.1%; Score 108; DB 9; Length 45;
Best Local Similarity 64.4%; Pred. No. 5.2e-05;
Matches 29; Conservative 1; Mismatches 5; Indels 10; Gaps 2;
QY 1 AKKYAKKEKA-AKKAYK-----KEAKAAEAERAAKEAYEA 35
Db 1 AKKYAKKAERAKKAYKAAEAKKAAYKAAEAKAAEAKAAEAYEA 45

RESULT 3
US-09-816-989A-6
; Sequence 6, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-6
Query Match 65.1%; Score 108; DB 9; Length 86;
Best Local Similarity 57.4%; Pred. No. 0.0001;
Matches 27; Conservative 2; Mismatches 6; Indels 12; Gaps 1;
QY 1 AKKYAKKEKA-AKKAYK-----KKEAKAAEAERAAKEAYEA 35
Db 40 AKKYAKKAERKYEAAAEAKYKAAEAKKAYKAAEAKAAEAKAAEAYEA 86

RESULT 4
US-09-816-989A-4
; Sequence 4, Application US/09816989A
; Patent No. US20020115103A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-4
Query Match 63.6%; Score 105.5; DB 9; Length 66;
Best Local Similarity 43.9%; Pred. No. 0.00014;
Matches 29; Conservative 2; Mismatches 4; Indels 31; Gaps 1;
QY 1 AKKYAKKEKA-AKKAYKKEA-----KAKAAEAERAAK 29
Db 1 AKKYAKKEKAYAKAKAAEAKAAKAAKAAKAAEAKYKAAEAKAAK 60
QY 30 EAYEA 35
Db 61 EAYEA 66

RESULT 5
US-09-816-989A-7
; Sequence 7, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-7
Query Match 63.3%; Score 105; DB 9; Length 109;
Best Local Similarity 62.2%; Pred. No. 0.00027;
Matches 28; Conservative 2; Mismatches 5; Indels 10; Gaps 2;
QY 1 AKKYAKKEKA-AKKAY-----KKEA-----KAKAAEAERAAEAYEA 35
Db 65 AKKYAKKAERKYEAAAEAKKAAEAKKAAEAKKAAEAKAAEAYEA 109

RESULT 6
US-09-816-989A-3
; Sequence 3, Application US/09816989A
```

Patent No. US20020115103A1

GENERAL INFORMATION:

APPLICANT: Gad, Alexander

APPLICANT: Lis, Doris

TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK

TITLE OF INVENTION: AND FOR THERAPEUTIC USE

FILE REFERENCE: 2609/60807-A-PCT-US

CURRENT APPLICATION NUMBER: US/09/816,989A

CURRENT FILING DATE: 2001-03-23

PRIOR APPLICATION NUMBER: 60/101,693

PRIOR FILING DATE: 1998-09-25

PRIOR APPLICATION NUMBER: PCT/US99/22402

PRIOR FILING DATE: 1999-09-24

NUMBER OF SEQ ID NOS: 7

SOFTWARE: Patentin version 3.1

SEQ ID NO 3

LENGTH: 56

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide

US-09-816-989A-3

Query Match 63.0%; Score 104.5; DB 9; Length 56;

Best Local Similarity 51.8%; Pred. No. 0.00016;

Matches 29; Conservative 4; Mismatches 2; Indels 21; Gaps 2;

QY 1 AKKYAKKEKA---AKKAYK-----EAKAKAAEAAAKAAEAA 35

Db 1 AKKYAKKEKAYAKKAEKAAKAEKAYKAAEAKKAEKAYKAAEAKKAAEAAEAA 56

RESULT 7

US-09-816-989A-5

Sequence 5, Application US/09816989A

Patent No. US20020115103A1

GENERAL INFORMATION:

APPLICANT: Gad, Alexander

APPLICANT: Lis, Doris

TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK

TITLE OF INVENTION: AND FOR THERAPEUTIC USE

FILE REFERENCE: 2609/60807-A-PCT-US

CURRENT APPLICATION NUMBER: US/09/816,989A

CURRENT FILING DATE: 2001-03-23

PRIOR APPLICATION NUMBER: 60/101,693

PRIOR FILING DATE: 1998-09-25

PRIOR APPLICATION NUMBER: PCT/US99/22402

PRIOR FILING DATE: 1999-09-24

NUMBER OF SEQ ID NOS: 7

SOFTWARE: Patentin version 3.1

SEQ ID NO 5

LENGTH: 77

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide

US-09-816-989A-5

Query Match 63.0%; Score 104.5; DB 9; Length 77;

Best Local Similarity 65.8%; Pred. No. 0.00022;

Matches 25; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

QY 1 AKKYAKKEKAACKAY---KKEAKAKAAEAAAKAAEAAEAA 35

Db 40 AKKYAKKAEKAEKAYEAAEAAEAKYKAEAKAAEAAEAAEAA 77

RESULT 8

US-10-282-122A-53988

Sequence 53988, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: Patentin version 3.1

SEQ ID NO 53988

LENGTH: 157

TYPE: PRT

ORGANISM: Corynebacterium diptheriae

US-10-282-122A-53988

Query Match 49.4%; Score 82; DB 15; Length 157;

Best Local Similarity 64.5%; Pred. No. 0.13;

Matches 20; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 KKYAKKEKAAKAYKKEAKAAEAAEAAEAA 32

Db 126 KKKAKEAAKAAEAAEAAEAAEAAEAAEAA 156

RESULT 9

US-10-282-122A-68109

Sequence 68109, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A





***This Page Blank (uspto)***

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	166	100.0	35	4	US-09-405-743A-1	Sequence 1, Appli
2	108	65.1	45	4	US-09-405-743A-2	Sequence 2, Appli
3	108	65.1	45	4	US-09-405-743A-6	Sequence 6, Appli
4	105.5	63.6	66	4	US-09-405-743A-4	Sequence 4, Appli
5	105	63.3	109	4	US-09-405-743A-7	Sequence 7, Appli
6	104.5	63.0	56	4	US-09-405-743A-3	Sequence 3, Appli
7	104.5	63.0	77	4	US-09-405-743A-5	Sequence 5, Appli
8	86.5	52.1	176	4	US-09-248-796A-18922	Sequence 18922, A
9	76	45.8	407	4	US-09-252-991A-29581	Sequence 29581, A
10	67.5	40.7	32	1	US-08-152-488-13	Sequence 13, Appl
11	67.5	40.7	32	1	US-08-303-025-15	Sequence 15, Appl
12	67.5	40.7	32	1	US-08-677-304-13	Sequence 13, Appl
13	67.5	40.7	32	2	US-08-436-703B-2	Sequence 2, Appli
14	67.5	40.7	33	2	US-08-303-025-16	Sequence 16, Appl
15	67.5	40.7	33	2	US-08-436-703B-4	Sequence 4, Appli
16	66.5	39.5	11	4	US-09-419-497-100	Sequence 100, App
17	65.5	39.5	214	3	US-09-041-889-27	Sequence 27, Appl
18	65.5	39.5	214	4	US-09-417-264-27	Sequence 27, Appl
19	65.5	39.5	469	4	US-09-489-039A-13565	Sequence 13565, A
20	65	39.2	223	3	US-09-095-855-201	Sequence 201, App
21	65	39.2	223	4	US-09-205-426-201	Sequence 201, App
22	65	39.2	364	4	US-09-107-532A-5044	Sequence 5044, Ap
23	64.5	38.9	204	4	US-08-529-055-21	Sequence 21, Appl
24	64.5	38.9	585	4	US-09-134-000C-3802	Sequence 3802, Ap
25	64.5	38.9	8991	4	US-08-714-741-32	Sequence 32, Appl
26	64	38.6	700	4	US-09-107-532A-5094	Sequence 5094, Ap
27	63	38.0	148	4	US-09-248-796A-26989	Sequence 26989, A

Query Match 65.1%; Score 108; DB 4; Length 45;  
Best Local Similarity 64.4%; Pred. No. 2.4e-06;  
Matches 29; Conservative 1; Mismatches 5; Indels 10; Gaps 2;

QY 1 AKKYAKKEKA--AKKAYK-----KEAKAKAAEAAAKEAAAYEA 35  
||||| ||| ||||| | | : ||||| |||||  
DB 1 AKKYAKKAKAEKAKKAYKAAEAKKAAKYEKAAEAKAAKEAAAYEA 45  
||||| ||| ||||| | | : ||||| |||||

RESULT 3  
US-09-405-743A-6  
; Sequence 6, Application US/09405743A  
; Patent No. 6514938  
; GENERAL INFORMATION:  
; APPLICANT: Yeda Research and Development Co., Ltd.  
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS  
; FILE REFERENCE: 60807-A  
; CURRENT APPLICATION NUMBER: US/09/405,743A  
; CURRENT FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 86  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
; OTHER INFORMATION: PEPTIDE  
US-09-405-743A-6

Query Match 65.1%; Score 108; DB 4; Length 86;  
Best Local Similarity 57.4%; Pred. No. 4.7e-06;  
Matches 27; Conservative 2; Mismatches 6; Indels 12; Gaps 1;

QY 1 AKKYAKKEKAACKAY-----KKEAKAKAAEAAAKEAAAYEA 35  
||||| ||| ||| | | : ||||| |||||  
DB 40 AKKYAKAAKAEKKEYYAAAEAKYKAAKAYKAAEAKAAKEAAAYEA 86  
||||| ||| ||| | | : ||||| |||||

RESULT 4  
US-09-405-743A-4  
; Sequence 4, Application US/09405743A  
; Patent No. 6514938  
; GENERAL INFORMATION:  
; APPLICANT: Yeda Research and Development Co., Ltd.  
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS  
; FILE REFERENCE: 60807-A  
; CURRENT APPLICATION NUMBER: US/09/405,743A  
; CURRENT FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 66  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
; OTHER INFORMATION: PEPTIDE  
US-09-405-743A-4

Query Match 63.6%; Score 105.5; DB 4; Length 66;  
Best Local Similarity 43.9%; Pred. No. 7e-06;  
Matches 29; Conservative 2; Mismatches 4; Indels 31; Gaps 1;

QY 1 AKKYAKKEKAACKAYKKEA-----KAKAAEAAAK 29  
||||| ||| ||| | | : ||||| |||||  
DB 1 AKKYAKKAYAKAKKAAEAKKAAKAAEAKKAAEAKYKAAEAAKAAK 60  
||||| ||| ||| | | : ||||| |||||

QY 30 EAAVEA 35  
|||||  
DB 61 EAAVEA 66  
|||||

RESULT 5  
US-09-405-743A-7  
; Sequence 7, Application US/09405743A  
; Patent No. 6514938  
; GENERAL INFORMATION:  
; APPLICANT: Yeda Research and Development Co., Ltd.  
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS  
; FILE REFERENCE: 60807-A  
; CURRENT APPLICATION NUMBER: US/09/405,743A  
; CURRENT FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 109  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
; OTHER INFORMATION: PEPTIDE  
US-09-405-743A-7

Query Match 63.3%; Score 105; DB 4; Length 109;  
Best Local Similarity 62.2%; Pred. No. 1.4e-05;  
Matches 28; Conservative 2; Mismatches 5; Indels 10; Gaps 2;

QY 1 AKKYAKKEKAACKAY-----KKEA---KAKAAEAAAKEAAAYEA 35  
||||| ||| ||| | | : ||||| |||||  
DB 65 AKKYAKAAKAEKKEYYAAAEAKKAAEAKKAYKAAEAKAAKEAAAYEA 109  
||||| ||| ||| | | : ||||| |||||

RESULT 6  
US-09-405-743A-3  
; Sequence 3, Application US/09405743A  
; Patent No. 6514938  
; GENERAL INFORMATION:  
; APPLICANT: Yeda Research and Development Co., Ltd.  
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS  
; FILE REFERENCE: 60807-A  
; CURRENT APPLICATION NUMBER: US/09/405,743A  
; CURRENT FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 56  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
; OTHER INFORMATION: PEPTIDE  
US-09-405-743A-3

Query Match 63.0%; Score 104.5; DB 4; Length 56;  
Best Local Similarity 51.8%; Pred. No. 7.8e-06;  
Matches 29; Conservative 4; Mismatches 2; Indels 21; Gaps 2;

QY 1 AKKYAKKEKA---AKKAYK-----EAKAKAAEAAAKEAAAYEA 35  
||||| ||| ||| | | : ||||| |||||  
DB 1 AKKYAKKAYAKKAAEAKKAAEAKKAAEAKKAYKAAEAKKAAKEAAAYEA 56  
||||| ||| ||| | | : ||||| |||||

RESULT 7  
US-09-405-743A-5  
; Sequence 5, Application US/09405743A  
; Patent No. 6514938  
; GENERAL INFORMATION:  
; APPLICANT: Yeda Research and Development Co., Ltd.  
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS  
; FILE REFERENCE: 60807-A  
; CURRENT APPLICATION NUMBER: US/09/405,743A  
; CURRENT FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1



```

Matches 23; Conservative 5; Mismatches 7; Indels 6; Gaps 2;
Qy 1 AKKYAK---KEKAACKYKK---EAKAKAAEAAKAAEAYEA 35
    ||| | : ||| | : ||| | : ||| | : ||| |
Db 215 AKTRADEAKKAAEDAKKAAEDAKKAAEAKKAAEAKKAAEA 255

RESULT 10
US-08-152-488-13
; Sequence 13, Application US/08152488
; Patent No. 5534619
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrew, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J, Rohm, Esq.
; STREET: 512 Springfield Avenue
; CITY: Cranford
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07016-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6; ASCII (DOS)Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/152,488
; FILING DATE: 12-NOV-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REGISTRATION NUMBER: 28,664
; REFERENCE/DOCKET NUMBER: RM-7MG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-276-3344
; TELEFAX: 908-276-5543
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
; PUBLICATION INFORMATION:
; AUTHORS: N/A
; TITLE: N/A
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993
; US-08-152-488-13

Query Match 40.7%; Score 67.5; DB 1; Length 32;
Best Local Similarity 61.3%; Pred. No. 0.097;
Matches 19; Conservative 2; Mismatches 9; Indels 1; Gaps 1;
Qy 1 AKKYAKK-EKAACKYKKKAAKAAEAAAKE 30
    ||| | : ||| | : ||| | : ||| |
Db 2 AKKAAKKAACKAAKKAACKAAKKAACKAAKKAACK 32

RESULT 11
US-08-303-025-15

```

/ Sequence 15, Application US/08303025  
/ Patent No. 5614494  
/ GENERAL INFORMATION:  
/ APPLICANT: Wakefield, Thomas W.  
/ APPLICANT: Andrews, Philip C.  
/ APPLICANT: Stanley, James C.  
/ TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND  
/ TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN  
/ TITLE OF INVENTION: ANTICOAGULATION REVERSAL  
/ NUMBER OF SEQUENCES: 16  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Benita J. Rohm, Esq.  
/ STREET: 150 West Jefferson, Suite 2500  
/ CITY: Detroit  
/ STATE: Michigan  
/ COUNTRY: United States of America  
/ ZIP: 48226-4415  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy diskette 3.5" 1.44mb  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: MS-DOS v.6.22  
/ SOFTWARE: WordPerfect 6.1; ASCII (DOS)Text  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/303,025  
/ FILING DATE: 08-SEPT-1994  
/ CLASSIFICATION: 514  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: PCT/US92/06829  
/ FILING DATE: 14-AUG-1992  
/ APPLICATION NUMBER: US 08/152,488  
/ FILING DATE: 12-NOV-1993  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Rohm, Benita J.  
/ REFERENCE/DOCKET NUMBER: 7WH-060548-00231  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 313-496-7622  
/ TELEFAX: 313-496-8454  
/ INFORMATION FOR SEQ ID NO: 15:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 32 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS: N/A  
/ TOPOLOGY: N/A  
/ MOLECULE TYPE: peptide  
/ ORIGINAL SOURCE:  
/ ORGANISM: N/A  
/ PUBLICATION INFORMATION:  
/ AUTHORS: N/A  
/ TITLE: N/A  
/ DOCUMENT NUMBER: PCT/US92/08069  
/ FILING DATE: 14-AUG-1993  
/ US-08-303-025-15

Query Match 40.7%; Score 67.5; DB 1; Length 32;  
Best Local Similarity 61.3%; Pred. No. 0.097;  
Matches 19; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 1 AKKYAKK-EKAAKAYKKEAKAKAAEAAAKE 30  
||| ||| : ||| ||| ||| ||| : ||| :  
Db 2 AKKAAKAKKAAKAAKAAKAAKAAKAAK 32

RESULT 12  
US-08-677-304-13  
/ Sequence 13, Application US/08677304  
/ Patent No. 5721212  
/ GENERAL INFORMATION:  
/ APPLICANT: Wakefield, Thomas W.  
/ APPLICANT: Andrews, Philip C.  
/ APPLICANT: Stanley, James C.  
/ TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND  
/ TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN  
/ TITLE OF INVENTION: ANTICOAGULATION REVERSAL

/ NUMBER OF SEQUENCES: 13  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Benita J. Rohm, Esq.  
/ STREET: 512 Springfield Avenue  
/ CITY: Cranford  
/ STATE: New Jersey  
/ COUNTRY: United States of America  
/ ZIP: 07016-1811  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: MS-DOS  
/ SOFTWARE: WordPerfect 6; ASCII (DOS)Text  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/677,304  
/ FILING DATE:  
/ CLASSIFICATION: 530  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/152,488  
/ FILING DATE: 12-NOV-1993  
/ APPLICATION NUMBER: PCT/US92/08069  
/ FILING DATE: 14-AUG-1993  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Rohm, Benita J.  
/ REGISTRATION NUMBER: 28,664  
/ REFERENCE/DOCKET NUMBER: RM-7WG  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: 908-276-3344  
/ TELEFAX: 908-276-5543  
/ INFORMATION FOR SEQ ID NO: 13:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 32 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS: No. 5721212 Relevant  
/ TOPOLOGY: No. 5721212 Relevant  
/ MOLECULE TYPE: peptide  
/ ORIGINAL SOURCE:  
/ ORGANISM: N/A  
/ PUBLICATION INFORMATION:  
/ AUTHORS: N/A  
/ TITLE: N/A  
/ PUBLICATION INFORMATION:  
/ DOCUMENT NUMBER: PCT/US92/08069  
/ FILING DATE: 14-AUG-1993  
/ US-08-677-304-13

Query Match 40.7%; Score 67.5; DB 1; Length 32;  
Best Local Similarity 61.3%; Pred. No. 0.097;  
Matches 19; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 1 AKKYAKK-EKAAKAYKKEAKAKAAEAAAKE 30  
||| ||| : ||| ||| ||| ||| : ||| :  
Db 2 AKKAAKAKKAAKAAKAAKAAKAAKAAK 32

RESULT 13  
US-08-436-703B-2  
/ Sequence 2, Application US/08436703B  
/ Patent No. 5919761  
/ GENERAL INFORMATION:  
/ APPLICANT: Wakefield, Thomas W.  
/ APPLICANT: Andrews, Philip C.  
/ APPLICANT: Stanley, James C.  
/ TITLE OF INVENTION: NOVEL PEPTIDES FOR  
/ TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR  
/ TITLE OF INVENTION: WEIGHT HEPARIN  
/ TITLE OF INVENTION: ANTICOAGULATION REVERSAL  
/ NUMBER OF SEQUENCES: 18  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Benita J. Rohm, Esq.  
/ STREET: 6601 Woodward Avenue  
/ CITY: Suite 1525  
/ CITY: Detroit

STATE: Michigan  
COUNTRY: United States of America  
ZIP: 48226  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk 1.44Mb, 3.5"  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 6;  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/436,703B  
FILING DATE: 08-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: N/A  
FILING DATE: N/A  
ATTORNEY/AGENT INFORMATION:  
NAME: Rohm, Benita J.  
REGISTRATION NUMBER: 28,664  
REFERENCE/DOCKET NUMBER: 7WK-060548-00233  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 313-965-1976  
TELEFAX: 313-965-1951  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: N/A  
TOPOLOGY: N/A  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: N/A  
PUBLICATION INFORMATION:  
AUTHORS: N/A  
TITLE: N/A

US-08-436-703B-2

Query Match 40.7%; Score 67.5; DB 2; Length 32;  
Best Local Similarity 61.3%; Pred. No. 0.097;  
Matches 19; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 1 AKYAKK-EKAACKAYKKEAKAKAEAAAKE 30  
||| ||| : ||||| | ||| ||| :  
DB 2 AKKAACKAKKAACKAKKAACKAKKAACK 32

RESULT 14  
US-08-303-025-16  
Sequence 16, Application US/08303025  
Patent No. 561494  
GENERAL INFORMATION:  
APPLICANT: Wakefield, Thomas W.  
APPLICANT: Andrews, Philip C.  
TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND  
TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN  
TITLE OF INVENTION: ANTICOAGULATION REVERSAL  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Benita J, Rohm, Esq.  
STREET: 150 West Jefferson, Suite 2500  
CITY: Detroit  
STATE: Michigan  
COUNTRY: United States of America  
ZIP: 48226-4415  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy diskette 3.5" 1.44Mb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS v.6.22  
SOFTWARE: WordPerfect 6.1; ASCII (DOS)Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/303,025  
FILING DATE: 08-SEPT-1994

CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/06829  
FILING DATE: 14-AUG-1992  
APPLICATION NUMBER: US 08/152,488  
FILING DATE: 12-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Rohm, Benita J.  
REFERENCE/DOCKET NUMBER: 7WH-060548-00231  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 313-496-7622  
TELEFAX: 313-496-8454  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 33 amino acids  
TYPE: amino acid  
STRANDEDNESS: N/A  
TOPOLOGY: N/A  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: N/A  
PUBLICATION INFORMATION:  
AUTHORS: N/A  
TITLE: N/A  
DOCUMENT NUMBER: PCT/US92/08069  
FILING DATE: 14-AUG-1993  
US-08-303-025-16

Query Match 40.7%; Score 67.5; DB 1; Length 33;  
Best Local Similarity 61.3%; Pred. No. 0.1;  
Matches 19; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 1 AKYAKK-EKAACKAYKKEAKAKAEAAAKE 30  
||| ||| : ||||| | ||| ||| :  
DB 3 AKKAACKAKKAACKAKKAACKAKKAACK 33

RESULT 15  
US-08-436-703B-4  
Sequence 4, Application US/08436703B  
Patent No. 5919761  
GENERAL INFORMATION:  
APPLICANT: Wakefield, Thomas W.  
APPLICANT: Andrews, Philip C.  
APPLICANT: Stanley, James C.  
TITLE OF INVENTION: NOVEL PEPTIDES FOR  
TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR  
TITLE OF INVENTION: WEIGHT HEPARIN  
TITLE OF INVENTION: ANTICOAGULATION REVERSAL  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Benita J, Rohm, Esq.  
STREET: 6601 Woodward Avenue  
STREET: Suite 1525  
CITY: Detroit  
STATE: Michigan  
COUNTRY: United States of America  
ZIP: 48226  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk 1.44Mb, 3.5"  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 6;  
SOFTWARE: ASCII (DOS)Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/436,703B  
FILING DATE: 08-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: N/A  
FILING DATE: N/A  
ATTORNEY/AGENT INFORMATION:  
NAME: Rohm, Benita J.

```

; REGISTRATION NUMBER: 28,664
; REFERENCE/DOCKET NUMBER: 7WK-060548-00233
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 313-965-1976
; TELEFAX: 313-965-1951
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
; PUBLICATION INFORMATION:
; AUTHORS: N/A
; TITLE: N/A
; US-08-436-703B-4

```

```

Query Match 40.7%; Score 67.5; DB 2; Length 33;
Best Local Similarity 61.3%; Pred. No. 0.1;
Matches 19; Conservative 9; Mismatches 9; Indels 1; Gaps 1;

```

```

Qy 1 AKYAKK-EKAAKAYKKEAKAAEAAKE 30
   ||| ||| : ||||| | ||| |||
Db 3 AKYAKKAKKAAKAAKAAKAAKAAKAAK 33

```

```

Search completed: December 14, 2004, 05:50:15
Job time : 10.9635 secs

```



Db 220 KKAEEKAAAEKAAADKAAAEKAAADKKAAA-AKAAAEKAAAEKAAAE 267

RESULT 3

G85576

membrane spanning protein Tola [imported] - Escherichia coli (strain O157:H7, substrain C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004

C;Accession: G85576

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; Fouts, D.E.; Keller, J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, L.; Grotbeck, E.J.; 2001

Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: G85576

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-394 <STO>

A;Cross-references: UNIPROT:Q8X965; GB:AE005174; NID:g12513672; PIDN:AGS5075.1; GSPDB:G

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: tola

Query Match 46.0%; Score 98; DB 2; Length 394;

Best Local Similarity 65.3%; Pred. No. 0.03;

Matches 32; Conservative 2; Mismatches 9; Indels 6; Gaps 3;

QY 2 KKYAKKAAEKAKA---KK--AYKAAEKAKAYEKAAAEKAAAEKAAAE 45

Db 220 KKAEEKAAAEKAAADKAAAEKAAADKKAAA-AKAAAEKAAAEKAAAE 267

RESULT 4

JV0057

tola protein - Escherichia coli (strain K-12)

C;Species: Escherichia coli

C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004

C;Accession: JV0057; B64810

R;Levengood, S.K.; Webster, R.E.

J. Bacteriol. 171, 6600-6609, 1989

A;Title: Nucleotide sequences of the tola and tolB genes and localization of their products

A;Reference number: JV0057; MUID:90078104; PMID:2687247

A;Accession: JV0057

A;Molecule type: DNA

A;Residues: 1-421 <LEV>

A;Cross-references: UNIPROT:P19934; GB:M28232; NID:g148018; PIDN:ABA24683.1; PID:g148019

A;Experimental source: strain JM105

A;Note: the authors translated the initiation codon GTG for residue 1 as Val

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: B64810

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-421 <BLAT>

A;Cross-references: GB:AE000177; GB:U00096; NID:g1786955; PIDN:AACT3833.1; PID:g1786960;

A;Experimental source: strain K-12, substrain MG1655

C;Comment: tola and tolB proteins are necessary for colicins E2, E3, A, and K to reach the cell surface

C;Genetics:

A;Gene: tola

A;Map position: 17 min

A;Start codon: GTG

C;Keywords: nucleotide binding; P-loop; transmembrane protein

F;14-34/Domain: transmembrane #status predicted <MSS>

F;78-301/Domain: helical #status predicted <HSS>

F;355-362/Region: nucleotide-binding motif A (P-loop)

Query Match 46.0%; Score 98; DB 2; Length 421;

Best Local Similarity 65.3%; Pred. No. 0.032;

Matches 32; Conservative 2; Mismatches 9; Indels 6; Gaps 3;

QY 2 KKYAKKAAEKAKA---KK--AYKAAEKAKAYEKAAAEKAAAEKAAAE 45

Db 247 KKAEEKAAAEKAAADKAAAEKAAADKKAAA-AKAAAEKAAAEKAAAE 294

RESULT 5

AEI317

hypothetical protein lmo1941 [imported] - Listeria monocytogenes (strain EGD-e)

C;Species: Listeria monocytogenes

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004

C;Accession: AEI317

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maiok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AEI317

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-239 <GLA>

A;Cross-references: UNIPROT:Q8Y5W4; GB:NC\_003210; PIDN:CAD00019.1; PID:g16411394; GSPDB:G

A;Experimental source: strain EGD-e

C;Genetics:

A;Gene: lmo1941

Query Match 44.4%; Score 94.5; DB 2; Length 239;

Best Local Similarity 56.2%; Pred. No. 0.043;

Matches 27; Conservative 7; Mismatches 9; Indels 5; Gaps 2;

QY 2 KKYAKKAAEKAKK---AYKAAEKAKK--AAKYEKAAAEKAAAEKAAAE 44

Db 124 KAAAEKAEADKKQBEDAVKAAANAKKEQEAEEKAAADKAAAEKAAAE 171

RESULT 6

AEI689

hypothetical protein homolog lin2055 [imported] - Listeria innocua (strain Clip11262)

C;Species: Listeria innocua

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004

C;Accession: AEI689

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maiok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AEI689

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-243 <GLA>

A;Cross-references: UNIPROT:Q92A67; GB:AL592022; PIDN:CAC97285.1; PID:g16414556; GSPDB:G

A;Experimental source: strain Clip11262

C;Genetics:

A;Gene: lin2055

Query Match 42.3%; Score 90; DB 2; Length 243;

Best Local Similarity 45.5%; Pred. No. 0.11;

Matches 30; Conservative 3; Mismatches 11; Indels 22; Gaps 2;

QY 1 AKKYA-KKAAEKAKAYKAAAEKAKA-----AKYEKAAAEKAAAE 38

Db 106 AKKAAEEKAAAEKAAAEKAAAEKAAADKKSQDEAAKAAAKKQEQAEEKAAAEKAAAE 165

QY 39 KEAAVE 44

Db 166 DKAAKE 171

[illegible]





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 50.5063 Seconds  
(without alignments)  
512.646 Million cell updates/sec

Title: US-10-792-311-2

Perfect score: 213

Sequence: 1 AKYAKAKAEKAKYAKAA.....AKYKAAAEKAAKAAEAYE 45

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	108.5	50.9	379	2	Q7Wfn5	bordetella
2	103.5	48.6	210	1	H1_LYTP1	P06144 lytechinus
3	103.5	48.6	373	2	Q7W477	bordetella
4	98.5	46.2	461	2	Q7UR44	Q7ur44 rhodospirillum rubrum
5	98	46.0	394	2	Q7AG18	Q7ag18 escherichia coli
6	98	46.0	394	2	Q8X965	Q8x965 escherichia coli
7	98	46.0	421	1	T0A4_ECOLI	P19934 escherichia coli
8	98	46.0	421	2	Q8FJTI	Q8fjti escherichia coli
9	95.5	44.8	198	2	Q6MIU4	Q6miu4 bdellovibrio bacteriovorus
10	95.5	44.8	198	2	CAE80819	Caes80819 bdellovibrio bacteriovorus
11	95	44.6	413	2	Q7C3Q4	Q7c3q4 shigella flexneri
12	95	44.6	413	2	Q83SA1	Q83sa1 shigella flexneri
13	94.5	44.4	239	2	Q8Y5W4	Q8y5w4 listeria monocytogenes
14	93.5	43.9	248	2	Q81H69	Q81h69 bacillus cereus
15	92	43.2	117	2	Q89DA0	Q89da0 bradyrhizobium elkanii
16	91	42.7	244	2	Q9AJX2	Q9ajx2 streptomyces
17	90.5	42.5	197	2	Q7W3X2	Q7w3x2 bordetella pertussis
18	90	42.3	243	2	Q92A67	Q92a67 listeria monocytogenes
19	90	42.3	280	2	Q6CEE5	Q6cee5 yarrowia lipolytica
20	90	42.3	1391	1	MS2_DROHY	Q08696 drosophila
21	89.5	42.0	208	1	DBH_MYCSM	Q92hc5 mycobacterium
22	89.5	42.0	211	1	H1B_STRPU	P15969 strongyloides
23	89.5	42.0	214	1	DBH_MYCBO	Q9xb18 mycobacterium
24	89.5	42.0	214	1	DBH_MYCTU	P95109 mycobacterium
25	89.5	42.0	244	2	Q71Y74	Q71y74 listeria monocytogenes
26	89.5	42.0	244	2	AAT04740	Aat04740 listeria monocytogenes
27	89.5	42.0	549	2	Q75CR3	Q75cr3 ashbya gossypii
28	89.5	42.0	549	2	AAS51084	Aas51084 ashbya gossypii
29	89	41.8	182	2	Q8RL36	Q8rl36 burkholderia
30	88.5	41.5	275	1	MS73_DROHY	Q01395 drosophila
31	88.5	41.5	372	2	Q9Wwx1	Q9wwx1 pseudomonas

32	88.5	41.5	372	2	Q88NI6	Q88ni6 pseudomonas
33	88	41.3	225	2	Q7WFA2	Q7wfa2 bordetella
34	88	41.3	420	2	Q73D27	Q73d27 bacillus cereus
35	88	41.3	420	2	AAS39818	Aas39818 bacillus cereus
36	88	41.3	899	2	Q9A5J6	Q9a5j6 caulobacter
37	87.5	41.1	1551	2	Q7SBU1	Q7sbu1 neurospora
38	86.5	40.6	200	2	Q8XVW7	Q8xv7 ralstonia
39	86.5	40.6	384	1	TMPB_TREPH	P29720 treponema
40	86	40.4	105	2	Q6N503	Q6n503 rhodospirillum rubrum
41	86	40.4	105	2	CAE28621	Caes28621 rhodospirillum rubrum
42	86	40.4	155	2	Q8PI40	Q8pi40 xanthomonas
43	86	40.4	168	1	RS16_COREF	Q8fp30 corynebacterium
44	86	40.4	198	1	HBHA_MYCTU	Q11142 mycobacterium
45	86	40.4	212	2	O93946	O93946 candida albicans

#### ALIGNMENTS

##### RESULT 1

ID	Q7Wfn5	PRELIMINARY;	PRT;	379 AA.
AC	Q7Wfn5;			
DT	01-OCT-2003 (Tremblrel. 25, Created)			
DT	01-OCT-2003 (Tremblrel. 25, Last sequence update)			
DE	01-MAR-2004 (Tremblrel. 26, Last annotation update)			
DE	Prolin-rich inner membrane protein.			
GN	OrderedLocusNames=BB4236;			
OS	Bordetella bronchiseptica (Alcaligenes bronchisepticus).			
OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;			
OC	Alcaligenaceae; Bordetella.			
OX	NCBI_TaxID=518;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=RB50 / ATCC BAA-588;			
RX	MEDLINE=2821954; PubMed=12910271; DOI=10.1038/ngi1227;			
RA	Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,			
RA	Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,			
RA	Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,			
RA	Achtman M., Atkin R., Baker S., Basham D., Bason N., Chervach I.,			
RA	Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,			
RA	Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,			
RA	Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,			
RA	Rabinowitz E., Rutter S., Sanders M., Saunders D., Seeger K.,			
RA	Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,			
RA	Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;			
RT	"Comparative analysis of the genome sequences of Bordetella pertussis,			
RT	Bordetella parapertussis and Bordetella bronchiseptica.";			
RL	Nat. Genet. 35:32-40(2003).			
DR	EMBL; BX640449; CAE34600.1; -			
DR	GO; GO:0016020; C:membrane; IEA.			
DR	GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.			
DR	GO; GO:0008565; F:protein transporter activity; IEA.			
DR	GO; GO:0015031; P:protein transport; IEA.			
DR	InterPro; IPR010528; ToIA.			
DR	InterPro; IPR006260; TonB_C.			
DR	Pfam; PF06519; ToIA; 1.			
DR	TIGRFAMs; TIGR01352; tonB_Cterm; 1.			
KW	Complete proteome.			
SQ	SEQUENCE 379 AA; 40776 MW; C657B5AAE97EBDD3 CRC64;			

Query Match 50.9%; Score 108.5; DB 2; Length 379;  
Best Local Similarity 65.1%; Pred. No. 0.0073;  
Matches 28; Conservative 5; Mismatches 7; Indels 3; Gaps 1;

QY 5 AKKAAEKKAKYKAAEKKAAKYE---KAAEKKAAEAYE 44

DB 201 AKKAAEKKAAEKKAAEKKAAEKKAAEKKAAEKKAAE 243

##### RESULT 2

HI\_LYTP1 STANDARD; PRT; 210 AA.

[illegible]

```

RESULT 5
Q7AGI8      PRELIMINARY;      PRT;      394 AA.
AC Q7AGI8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Membrane spanning protein TolA.
GN OrderedLocusNames=ECs0774;
OS Escherichia coli O157:H7;
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
DR EMBL; AP002553; BAB34197.1; -.
DR InterPro; IPR010528; TolA.
DR Pfam; PF06519; TolA; 1.
SQ SEQUENCE 394 AA; 40517 MW; 5B58D8B8230BDE28 CRC64;

Query Match      46.0%; Score 98; DB 2; Length 394;
Best Local Similarity 65.3%; Pred. No. 0.081;
Matches 32; Conservative 2; Mismatches 9; Indels 6; Gaps 3;

QY      2 KKYAKKAKEA---KK--AYKAAEAKKAAYEKAAAEKAAAEKAAAEYA 45
DB      220 KKAEEKAAAEKAAADKKAEEKAAADKKAEEKAAAEKAAAEKAAAE 267

RESULT 6
Q8X965      PRELIMINARY;      PRT;      394 AA.
AC Q8X965;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 26, Last annotation update)
DE Membrane spanning protein, required for outer membrane integrity.
GN Name=tolA; OrderedLocusNames=z0907;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
DR EMBL; AE005252; AAG55075.1; -.
DR PIR; F90725; F90725.
DR PIR; G85576; G85576.
DR HSSP; P19934; ITOL.
DR InterPro; IPR010528; TolA.
DR Pfam; PF06519; TolA; 1.
KW Complete proteome.
SQ SEQUENCE 394 AA; 40517 MW; 5B58D8B8230BDE28 CRC64;

Query Match      46.0%; Score 98; DB 2; Length 394;
Best Local Similarity 65.3%; Pred. No. 0.081;
Matches 32; Conservative 2; Mismatches 9; Indels 6; Gaps 3;

QY      2 KKYAKKAKEA---KK--AYKAAEAKKAAYEKAAAEKAAAEKAAAEYA 45
DB      220 KKAEEKAAAEKAAADKKAEEKAAADKKAEEKAAAEKAAAEKAAAE 267

RESULT 7
TOLA_ECOLI
ID TOLA_ECOLI      STANDARD;      PRT;      421 AA.
AC P19934;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE TolA protein.
GN Name=tolA; Synonyms=cim, excC, lky; OrderedLocusNames=b0739;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / JM105;
RX MEDLINE=90078104; PubMed=2687247;
RA Levgood S.K., Webster R.E.;
RT "Nucleotide sequences of the tolA and tolB genes and localization of
RT their products, components of a multistep translocation system in
RT Escherichia coli.";
RL J. Bacteriol. 171:6600-6609(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Colliado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [4]
RP DOMAINS.
RX MEDLINE=91296736; PubMed=2068069;
RA Levgood S.K., Beyer W.F. Jr., Webster R.E.;
RT "TolA: a membrane protein involved in colicin uptake contains an
RT extended helical region.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5939-5943(1991).
RN [5]
RP INTERACTION WITH PORINS.
RX MEDLINE=97133271; PubMed=8978668;
RA Derouiche R., Gavioli M., Benedetti H., Prilipov A., Lazdunski C.,
RA Lloubes R.;
RT "TolA central domain interacts with Escherichia coli porins.";
RL EMBO J. 15:6408-6415(1996).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 298-421.
RX MEDLINE=99332679; PubMed=10404600;
RA Lubkowski J., Hennecke F., Plueckhuhn A., Wlodawer A.;
RT "Filamentous phage infection: crystal structure of g3p in complex with
RT its coreceptor, the C-terminal domain of TolA.";

```



**D<sub>b</sub>**     | | |||| | : | | | |||| : ||| : ||| : ||| |  
32 AAKPAKATATAKAAPRAKKAAPKAAAPKAAAPKAAVTKAAPRA 77

## RESIT.T 10

CA80819	PRELIMINARY; PRT; 198 AA.
ID	CAE80819
AC	CAE80819;
DT	02-MAR-2004 (TrEMBLrel. 27, Created)
DT	02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT	02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE	Hypothetical protein.
GN	BD3054.
OS	Bdellovibrio bacteriovorus.
OC	Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC	Bdellovibrionaceae; Bdellovibrio.
OX	NCBI_TaxID=959;
[1]	SEQUENCE FROM N. A.
RN	STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RC	PubMed=14752164;
RX	Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA	Keller H., Lambert C., Evans K.J., Goesmann A., Meyer F.,
RA	Sockett R.E., Schuster S.C.;
RT	"A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
RT	genomic perspective.";
RL	Science 303:689-692(2004).
RL	EMBL; BX842654; CAE80819.1; -.
DR	Hypothetical protein.
KW	HYPOTHETICAL
SQ	SEQUENCE 198 AA; 198075 MW; 9942AF2CD8D65D33 CRC64;

Query Match	44.8%	Score	95.5	DB	2	Length	198
Best Local Similarity	58.7%	Pred. No.	0.076				
Matches	27	Conservative	4	Mismatches	14	Indels	1
						Gaps	1

Qy 1 AKYAKKAEKAKAYKAAEAKKAK-YEKAAAEKAAAEKAAEYEA 45  
| | | | | : | | | | | : | | | | |  
Db 32 AAKPAKATAKKAAKPAKKAAPKKAAPKKAAPKKAAPKKAAPKKAAPK 77  
| | | | | : | | | | | : | | | | |

## RESULT 11

Q80C04		PRELIMINARY;	PRT; 413 AA.
ID	Q7C2Q4		
AC	Q7C2Q4;		
DT	05-JUL-2004 (TrEMBLrel. 27, Created)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)		
DE	Membrane spanning protein.		
GN	Name=toIA; OrderedLocusNames=S0571;		
OS	Shigella flexneri.		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		
OC	Enterobacteriaceae; Shigella.		
OX	NCBI_TaxID=623;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=2457T;		
RX	MEDLINE=22590274; PubMed=13704152;		
RA	Weil J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,		
RA	Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,		
RA	Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,		
RA	Schwartz D.C., Blattner F.R.;		
RT	"Complete genome sequence and comparative genomics of Shigella		
RT	flexneri serotype 2a strain 2457T.";		
RL	Infect. Immun. 71:2775-2786(2003).		
DR	ENML; AE016979; AAP16075.1; -.		
DR	InterPro; IPR010528; ToIA.		
DR	Pfam; PF06519; ToIA; 1.		
SQ	SEQUENCE 413 AA; 42355 MW; 93E10F2C5DB60DE8 CRC64;		

Query Match	44.6%	Score 95;	DB 2;	Length 413;
Best Local Similarity	63.3%	Pred. No. 0.17;		
Matches 31;	Conservative	3;	Mismatches 9;	Indels 6;
Matches 31;	Conservative	3;	Mismatches 9;	Indels 6;
Matches 31;	Conservative	3;	Mismatches 9;	Indels 6;

Qy 2 KKYAKKAAEKA---KK--AYKAAEA KKA KYE KAAAEKAAAEKAAEYEA 45

D6 239 KKAEEKAAADKAAADKKAEEKAAADKAAA-AKAAAEKAAAKAKAAAE 286

## RESULT 12

Q83SA1	Q83SA1	PRELIMINARY;	PRT;	413 AA.
ID	Q83SA1			
AC	Q83SA1;			
DT	01-JUN-2003 (TrEMBLrel. 24, Created)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	Membrane spanning protein, required for outer membrane integrity.			
GN	Name=tolA; OrderedLocusNames=SF0558;			
OS	Shigella flexneri.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Shigella.			
OX	NCBI_TaxID=623;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=301 / Serotype 2a;			
RX	MEDLINE=22272406; PubMed=12384590;			
RA	Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,			
RA	Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,			
RA	Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,			
RA	Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,			
RA	Yu J.;			
RT	"Genome sequence of Shigella flexneri 2a: insights into pathogenicity			
RT	through comparison with genomes of Escherichia coli K12 and O157.";			
RL	Nucleic Acids Res. 30:4432-4441(2002).			
DR	ENBL; AE015086; AAN42202.1; --			
DR	HSSP; P15934; 1TOL.			
DR	InterPro; IPR010528; Tola.			
DR	Pfam; PF06519; Tola; 1.			
KW	Complete proteome.			
SO	SEQUENCE 413 AA; 42355 MW; 93E10F2C5DE60DE8 CRC64;			

Query Match 44.6%; Score 95; DB 2; Length 413;  
Best Local Similarity 63.3%; Pred. No. 0.17;  
Matches 31; Conservative 3; Mismatches 9; Indels 6; Gaps 3;

**QY** 2 KKYAKKAEKA ---KK--AYKAAEAKKAAKYEKAAAEEKAAAKEAAYEA 45  
.:|:|:|:|:| | | | | | | | | | :| |  
**pB** 219 KKAEEKAAADKAAADKKAEEKAAADKKAEE---AKAAAEKAAAKAKAAAEA 286

## RESIT.T 13

```

AC AC Q8Y5W4 PRELIMINARY; PRT; 239 AA.
AC Q8Y5W4;
AC Q8Y5W4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lmo1941 protein.
DE OrderedLocustNames=lmo1941;
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
ON NCBI_TaxID=1639;
RX [1]
RX SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RC MEDLINE=1537279; PubMed=11679669;
RA Glaser P., Frangoul L., Bloecher H. C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Buecher H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deboux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fiehi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krefz J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Mailcounam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsieck G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.,

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 49.7468 Seconds  
(without alignments)  
324.499 Million cell updates/sec

Title: US-10-792-311-2  
Perfect score: 213  
Sequence: 1 AKYAKKAAKAAKAYKAA.....AKYKAAAKAAKAAKAAAYEA 45

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003s:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	213	100.0	45	3	AAy82572 Copolymer
2	138	64.8	109	3	AAy82577 Copolymer
3	134.5	63.1	56	3	AAy82573 Copolymer
4	131	61.5	77	3	AAy82575 Copolymer
5	126.5	59.4	86	3	AAy82576 Copolymer
6	120.5	56.6	66	3	AAy82574 Copolymer
7	108.5	50.9	428	6	ABU27824 Protein e
8	108	50.7	35	3	AAy82571 Copolymer
9	99.5	46.7	106	2	AAR06446 Recombina
10	98	46.0	46	2	AAR28871 High affi
11	98	46.0	421	6	ABU28559 Protein e
12	95	44.6	323	6	ABU31397 Protein e
13	95	44.6	469	7	ABO67048 Klebsiell
14	94.5	44.4	239	5	ABB49123 Listeria
15	94.5	44.4	239	6	ABU32619 Protein e
16	94	44.1	154	2	AAR06445 Recombina
17	91	42.7	223	2	AAy14928 Amino aci
18	91	42.7	223	6	ABP70903 Mycobacte
19	89.5	42.0	214	2	AAy34055 M. tuberc
20	89.5	42.0	214	2	AAy57353 M. tuberc
21	89.5	42.0	214	6	ABU34623 Protein e
22	89.5	42.0	214	6	ABU36893 Protein e
23	88.5	41.5	372	6	ABU40185 Protein e
24	87	40.8	334	4	ABG28693 Novel hum
25	86	40.4	39	2	AAW44934 Mycobacte

26	86	40.4	39	7	ADP45561	Adf45561 Methylate
27	86	40.4	198	2	AAW43082	Aaw43082 Mycobacte
28	86	40.4	198	2	AAW44936	Aaw44936 Mycobacte
29	86	40.4	347	6	ABU38313	Abu38313 Protein e
30	86	40.4	347	6	ABJ18771	Abj18771 Pseudomon
31	86	40.4	407	7	ABO80835	AbO80835 Pseudomon
32	85.5	40.1	205	3	AAE20575	Aab20575 Mycobacte
33	85	39.9	427	4	AAg70868	Aag70868 C albican
34	85	39.9	427	8	ADP99051	Adp99051 C. albica
35	83	39.0	389	6	ABU39221	Abu39221 Protein e
36	82	38.5	165	4	AAG91997	Aag91997 C glutami
37	81.5	38.3	357	6	ABM67869	ABm67869 Photorhab
38	81.5	38.3	497	7	ADH88105	Adh88105 Enterococ
39	81	38.0	80	5	ABG71044	Abg71044 Tumour ne
40	79.5	37.3	407	6	ABU47123	Abu47123 Protein e
41	79	37.1	212	6	ABU33811	Abu33811 Protein e
42	78.5	36.9	875	6	ABU22879	Abu22879 Protein e
43	78	36.6	157	6	ABU26064	Abu26064 Protein e
44	78	36.6	309	7	ABO23523	AbO23523 Pseudomon
45	78	36.6	316	7	ABO84211	ABO84211 Pseudomon

ALIGNMENTS

RESULT 1  
AAy82572 standard; peptide; 45 AA.  
XX AAy82572;

AC AAy82572;  
XX  
DT 28-JUL-2000 (first entry)  
XX  
DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:2.  
XX  
KW Copolymer; molecular weight marker; TV-marker; immune disease;  
KW glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;  
KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;  
KW antididiabetic; thymostimetic; haemostatic; antipsoriatic; dermatological;  
KW antianemic; immunosuppressive; demyelinating disease; osteoarthritis;  
KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;  
KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;  
KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;  
KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;  
KW pemphigus vulgaris; systemic lupus erythematosus.  
XX  
OS Unidentified.

XX WO200018794-A1.  
XX  
PD 06-APR-2000.  
XX  
PF 24-SEP-1999; 99WO-US022402..  
XX  
PR 25-SEP-1998; 98US-0101693P.  
XX  
PA (YEDA ) YEDA RES & DEV CO LTD.  
PA (TEVA-) TEVA PHARM USA INC.  
XX  
PI Gad A, Lis D;  
XX  
DR WPI; 2000-317499/27.  
XX  
PT Copolymer 1 related polypeptides used as molecular weight markers for  
PT glatiramer acetate and prevention of immune diseases.  
XX  
PS Claim 10; Page 14; 72pp; English.

XX  
CC AAy82571 to AAy82577 represent specifically claimed copolymer molecular  
CC weight TV-marker polypeptides from the present invention. The present  
CC invention describes polypeptides (I) for determining the molecular weight  
CC of a copolymer (CP), which has an identified molecular weight and an  
CC amino acid composition corresponding to the copolymer. The polypeptides





```

XX PS Claim 10; Page 14; 72pp; English.
XX CC AAY82571 to AAY82577 represent specifically claimed copolymer molecular
XX CC weight TV-marker polypeptides from the present invention. The present
XX CC invention describes polypeptides (I) for determining the molecular weight
XX CC of a copolymer (CP), which has an identified molecular weight and an
XX CC amino acid composition corresponding to the copolymer. The polypeptides
XX CC of the invention are used as molecular weight markers for glatiramer
XX CC acetate related tetrapolymers. The polypeptides may also be used for
XX CC treating and preventing immune diseases in a mammal. Autoimmune diseases
XX CC which may be treated include either cell-mediated or antibody-mediated
XX CC diseases. Such diseases include arthritic conditions, demyelinating
XX CC arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune
XX CC disease, chronic immune thrombocytopaenia purpura, colitis, contact
XX CC sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's
XX CC syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,
XX CC psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-
XX CC graft-versus-host disease, and delayed-type hypersensitivity. The
XX CC polypeptides of the invention have defined molecular weights and physical
XX CC properties which are analogous to glatiramer acetate molecules, which
XX CC makes them ideal for use as molecular weight markers
XX CC
XX CC Sequence 56 AA;
XX CC
XX CC Query Match 63.1%; Score 134.5; DB 3; Length 56;
XX CC Best Local Similarity 68.4%; Pred. No. 1.4e-07;
XX CC Matches 39; Conservative 0; Mismatches 5; Indels 13; Gaps 4;
XX CC
QY 1 AKKYAKK-----AKAEKA-----KKAYKAAEAKK--AAKYKAAEKAKEAAEAAVEA 45
DB 1 AKKYAKKEKAYAKKAEKAAKAEKAAKAYKAAEAKKAEAKY-KAEAKAAKAEAAVEA 56

RESULT 4
AAY82575
ID AAY82575 standard; peptide; 77 AA.
AC AAY82575;
XX
XX 28-JUL-2000 (first entry)
XX
XX Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:5.
XX
XX Copolymer; molecular weight marker; TV-marker; immune disease;
XX glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;
XX osteopathic; immunosuppressive; antithyroid; antiinflammatory;
XX antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;
XX antianaemic; immunosuppressive; demyelinating disease; osteoarthritis;
XX inflammatory condition; multiple sclerosis; rheumatoid arthritis;
XX Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;
XX Hashimoto's disease; Graves disease; Guillain-Barre's syndrome; psoriasis;
XX pemphigus vulgaris; systemic lupus erythematosus.
XX
XX Unidentified.
XX
XX WO200018794-A1.
XX
XX 06-APR-2000.
XX
XX 24-SEP-1999; 99WO-US022402.
XX
XX 25-SEP-1998; 98US-0101693P.
XX
XX (YEDA ) YEDA RES & DEV CO LTD.
XX (TEVA-) TEVA PHARM USA INC.
XX
XX Gad A, Lis D;
XX

```

```

DR WPI; 2000-317499/27.
XX
XX Copolymer 1 related polypeptides used as molecular weight markers for
XX glatiramer acetate and for treatment and prevention of immune diseases.
XX
XX Claim 10; Page 14; 72pp; English.
XX
XX AAY82571 to AAY82577 represent specifically claimed copolymer molecular
XX weight TV-marker polypeptides from the present invention. The present
XX invention describes polypeptides (I) for determining the molecular weight
XX of a copolymer (CP), which has an identified molecular weight and an
XX amino acid composition corresponding to the copolymer. The polypeptides
XX of the invention are used as molecular weight markers for glatiramer
XX acetate related tetrapolymers. The polypeptides may also be used for
XX treating and preventing immune diseases in a mammal. Autoimmune diseases
XX which may be treated include either cell-mediated or antibody-mediated
XX diseases. Such diseases include arthritic conditions, demyelinating
XX diseases and inflammatory conditions, e.g. multiple sclerosis, rheumatoid
XX arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune
XX oophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's
XX disease, chronic immune thrombocytopaenia purpura, colitis, contact
XX sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's
XX syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,
XX psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-
XX graft-versus-host disease, and delayed-type hypersensitivity. The
XX polypeptides of the invention have defined molecular weights and physical
XX properties which are analogous to glatiramer acetate molecules, which
XX makes them ideal for use as molecular weight markers
XX
XX Sequence 77 AA;
XX
XX Query Match 61.5%; Score 131; DB 3; Length 77;
XX Best Local Similarity 50.6%; Pred. No. 4.6e-07;
XX Matches 39; Conservative 0; Mismatches 6; Indels 32; Gaps 3;
XX
QY 1 AKKYAKK-----AKAEKA-----KKAYKAAEAKKAAKYE----- 29
DB 1 AKKYAKKEKAYAKKAEKAAKAEKAAKAYKAAEAKKAEAKKAYKAAEAKKAEAKKAYKAAEAKKAEAKK 60
QY 30 -KAAAEKAAEAAVEA 45
DB 61 YKAEAAKAAEAAVEA 77

RESULT 5
AAY82576
ID AAY82576 standard; peptide; 86 AA.
AC AAY82576;
XX
XX 28-JUL-2000 (first entry)
XX
XX Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:6.
XX
XX Copolymer; molecular weight marker; TV-marker; immune disease;
XX glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;
XX osteopathic; immunosuppressive; antithyroid; antiinflammatory;
XX antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;
XX antianaemic; immunosuppressive; demyelinating disease; osteoarthritis;
XX inflammatory condition; multiple sclerosis; rheumatoid arthritis;
XX Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;
XX Hashimoto's disease; Graves disease; Guillain-Barre's syndrome; psoriasis;
XX Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;
XX pemphigus vulgaris; systemic lupus erythematosus.
XX
XX Unidentified.
XX
XX WO200018794-A1.
XX
XX 06-APR-2000.
XX
XX 24-SEP-1999; 99WO-US022402.
XX

```





ID AAR06446 standard; protein; 106 AA.  
 AC AAR06446;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 03-JAN-1991 (first entry)  
 XX  
 DE Recombinant copolymer 1-19, myelin basic protein analogue.  
 XX  
 KW Recombinant copolymer 1; COP-1-19; myelin basic protein; MBP;  
 KW immunological activity; autoimmune encephalomyelitis; multiple sclerosis.  
 XX  
 OS Synthetic.  
 OS  
 PN BP383620-A.  
 XX  
 PD 22-AUG-1990.  
 XX  
 PF 16-FEB-1990; 90EP-00301700.  
 XX  
 PR 17-FEB-1989; 89US-00312541.  
 PR 07-FEB-1990; 90US-00473845.  
 XX  
 PA (REPK ) REPLIGEN CORP.  
 XX  
 PI Cook KS;  
 XX  
 DR WPI; 1990-255848/34.  
 DR N-PSDB; AAQ06446.  
 XX  
 PT Producing genes encoding random polymers of aminoacid(s) - for producing  
 PT recombinant polypeptide(s) with biological and/or immunological activity.  
 PS Disclosure; Fig 12; 25pp; English.  
 XX  
 CC To improve the expression of rCOP-1 polypeptides in E. coli, genes coding  
 CC for rCOP-1-19 were subcloned from pREV 2.1 to pBG3-2deltaN (deposit: 20-  
 CC NOV-1984 US4691009, NRRL B-15910), a plasmid used to express Protein A.  
 CC The resulting plasmids encode fusion proteins consisting of beta-  
 CC glucuronidase, Protein A, and rCOP-1 sequences. A methionine residue  
 CC occurs between the Protein A and rCOP-1 sequences, originating from the  
 CC 5' linker sequence, in order that the COP-1 polypeptide may be cleaved  
 CC from the fusion protein. rCOP-1-19 contains oligonucleotide duplexes  
 CC encoding the following segments: YKK, AAE, KAK, EKA, KKA, YEA, AKA, KEA,  
 CC and AAA. The N-terminal alanine residue is left behind following CNBr  
 CC cleavage of the fusion protein. The product prevents or arrests  
 CC experimental autoimmune encephalomyelitis. They are used to prevent,  
 CC arrest or control a demyelinating disorder, e.g. multiple sclerosis. They  
 CC may also be used as additives to hair care products to confer beneficial  
 CC effects on damaged hair or as supplements for diets deficient in certain  
 CC amino acids. See also AAQ05664. (Updated on 25-MAR-2003 to correct PA  
 CC field.)  
 XX  
 SQ Sequence 106 AA;  
 Query Match 46.7%; Score 99.5; DB 2; Length 106;  
 Best Local Similarity 58.3%; Pred. No. 0.0017;  
 Matches 28; Conservative 4; Mismatches 7; Indels 9; Gaps 2;  
 OY 3 KYAKKA-KAKAKAKYKAAEAKKAAKAAKAAKAA-----AKEA 41  
 | | | | : | | | : | : | | | | | | | | | | | | | | | |  
 Db 8 KAAKAYAEAKAKAKYKAAEAKKAAKAAEAKKAAEAKKAAKAAKAAE 55  
 RESULT 10  
 AAR28871  
 ID AAR28871 standard; peptide; 46 AA.  
 XX  
 AC AAR28871;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 23-MAR-1993 (first entry)  
 XX  
 DE High affinity macrophage mannose receptor ligand compound #9.  
 XX  
 KW glycopeptide; mannose; mannosylated; glycosylated; mannose receptor;  
 KW macrophages; monocytes; destroy; cytotoxicity; label; image; alter;  
 KW macrophage processing of antigen; MHC restriction; inflammation;  
 KW inflammatory diseases; macrophage secretory products; Crohn's disease;  
 KW Legionnaires disease; mononuclear phagocytes; HIV; AIDS;  
 KW lysosomal storage diseases; Gaucher's disease; asthma;  
 KW alveolar macrophages metastasis; systemic macrophages; deliver;  
 KW antigenic peptides; prevent transplant rejection; organ transplantation;  
 KW antitumour agents; cancer; toxins.  
 XX  
 OS Synthetic.  
 OS  
 PH Key  
 FT Modified-site 1  
 /note= "opt may have mannose, fucose, glucose or N-Ac-  
 FT glucosamine. May also have non interfering substituents."  
 FT 4  
 Modified-site 4  
 /note= "opt may have mannose, fucose, glucose or N-Ac-  
 FT glucosamine."  
 FT 7  
 Modified-site 7  
 /note= "opt may have mannose, fucose, glucose or N-Ac-  
 FT glucosamine."  
 FT 10  
 Modified-site 10  
 /note= "opt may have mannose, fucose, glucose or N-Ac-  
 FT glucosamine."  
 FT 13  
 Modified-site 13  
 /note= "opt may have mannose, fucose, glucose or N-Ac-  
 FT glucosamine."  
 FT 16  
 Modified-site 16  
 /note= "opt may have mannose, fucose, glucose or N-Ac-  
 FT glucosamine."  
 FT 19  
 Modified-site 19  
 /note= "opt may have mannose, fucose, glucose or N-Ac-  
 FT glucosamine."  
 FT 22  
 Modified-site 22  
 /note= "opt may have mannose, fucose, glucose or N-Ac-  
 FT glucosamine."  
 FT 25  
 Modified-site 25  
 /note= "opt may have mannose, fucose, glucose or N-Ac-  
 FT glucosamine."  
 FT 28  
 Modified-site 28  
 /note= "opt may have mannose, fucose, glucose or N-Ac-  
 FT glucosamine."  
 FT 31  
 Modified-site 31  
 /note= "opt may have mannose, fucose, glucose or N-Ac-  
 FT glucosamine."  
 FT 34  
 Modified-site 34  
 /note= "opt may have mannose, fucose, glucose or N-Ac-  
 FT glucosamine."  
 FT 37  
 Modified-site 37  
 /note= "opt may have mannose, fucose, glucose or N-Ac-  
 FT glucosamine."  
 FT 40  
 Modified-site 40  
 /note= "opt may have mannose, fucose, glucose or N-Ac-  
 FT glucosamine."  
 FT 43  
 Modified-site 43  
 /note= "opt may have mannose, fucose, glucose or N-Ac-  
 FT glucosamine."  
 FT 46  
 Modified-site 46  
 /note= "opt may have mannose, fucose, glucose or N-Ac-  
 FT glucosamine. May also have non interfering substituents."  
 FT  
 XX  
 PN WO9219248-A1.  
 XX  
 PD 12-NOV-1992.  
 XX  
 PF 01-MAY-1992; 92WO-US003609.  
 XX  
 PR 03-MAY-1991; 91US-00694983.  
 PA (UNIW ) UNIV WASHINGTON.



PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 XX WPI; 2003-029926/02.  
 DR N-PSDB; ACA35267.  
 XX  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 XX Claim 25; SEQ ID NO 59321; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 323 AA;  
  
 Query Match 44.6%; Score 95; DB 6; Length 323;  
 Best Local Similarity 52.9%; Pred. No. 0.016;  
 Matches 27; Conservative 4; Mismatches 14; Indels 6; Gaps 1;  
  
 QY 1 AKKYAKAKAEKAKYKAAEAK-----KAAKYKAAAEKAAKAAEAAVEA 45  
 Db 92 AAEAQKAAEAAAKKQAEKKAQAEAKQAAEAKAAEAKAAEAKAAQKA 142  
  
 RESULT 13  
 ABO67048  
 ID ABO67048 standard; protein; 469 AA.  
 XX  
 AC ABO67048;  
 XX  
 XX 29-JUL-2004 (first entry)  
 DE Klebsiella pneumoniae polypeptide seqid 13565.  
 XX Recombinant expression vector; transcription regulatory element;  
 KW Klebsiella pneumoniae protein; antibacterial; Vaccine.

XX Klebsiella pneumoniae.  
 OS US6610836-B1.  
 PN 26-AUG-2003.  
 XX  
 PD 27-JAN-2000; 2000US-00489039.  
 XX  
 PF 29-JAN-1999; 99US-0117747P.  
 XX  
 PR (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PA Breton GL, Osborne M;  
 XX WPI; 2003-895346/82.  
 DR N-PSDB; ABD00619.  
 XX  
 XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for  
 PT preparing a vaccine composition against Klebsiella pneumoniae.  
 XX  
 PS Disclosure; SEQ ID NO 13565; 932pp; English.  
 XX  
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella  
 CC pneumoniae polypeptide. Also described are: a recombinant expression  
 CC vector comprising the nucleic acid, operably linked to a transcription  
 CC regulatory element; and a cell comprising the recombinant expression  
 CC vector. The nucleic acid is useful for preparing a vaccine composition  
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a  
 CC Klebsiella pneumoniae polypeptide of the invention  
 XX  
 SQ Sequence 469 AA;  
  
 Query Match 44.6%; Score 95; DB 7; Length 469;  
 Best Local Similarity 52.9%; Pred. No. 0.024;  
 Matches 27; Conservative 4; Mismatches 14; Indels 6; Gaps 1;  
  
 QY 1 AKKYAKAKAEKAKYKAAEAK-----KAAKYKAAAEKAAEAAVEA 45  
 Db 242 AAEAQKAAEAAAKKQAEKKAQAEAKQAAEAKAAEAKAAEAKAAQKA 292  
  
 RESULT 14  
 ABB49123  
 ID ABB49123 standard; protein; 239 AA.  
 XX  
 AC ABB49123;  
 XX  
 DT 05-FEB-2002 (first entry)  
 XX  
 DE Listeria monocytogenes protein #1827.  
 XX  
 KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;  
 KW vitamin B12; bacterial infection; disease.  
 XX  
 OS Listeria monocytogenes.  
 XX  
 XX WO200177335-A2.  
 PN 18-OCT-2001.  
 XX  
 PD 11-APR-2001; 2001WO-FR001118.  
 XX  
 PF 11-APR-2000; 2000FR-00004629.  
 PR (INSP ) INST PASTEUR.  
 XX  
 PA Buchrieser C, Frangeul L, Couve E, Rusniok C, Feihi H, Dehoux P;  
 PI Duseurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;  
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;  
 PI Dominguez-Bernal G, Garrido-Garcia P, Pierrez-Martinez A, Amend A;  
 PI Chakraborty T, Dommann E, Hain T, Berche P, Charbit A, Durant L;  
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;

Search completed: December 14, 2004, 06:01:11  
Job time : 52.7468 secs

**This Page Blank (uspto)**





```
US-09-816-989A-7
; Sequence 7, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816, 989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-7
Query Match 64.8%; Score 138; DB 9; Length 109;
Best Local Similarity 76.6%; Pred. No. 1.2e-07;
Matches 36; Conservative 0; Mismatches 7; Indels 4; Gaps 2;
QY 1 AKKYAKKAAKAAKAYKAAEAKK--AAKYKAAAEKAAAEKAAEAYEA 45
Db 65 AKKYAKAAKAE--KKEYYAAAEAKKAAKAYKAAEAKKAAEAYEA 109

RESULT 3
US-09-816-989A-3
; Sequence 3, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816, 989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-3
Query Match 63.1%; Score 134.5; DB 9; Length 56;
Best Local Similarity 68.4%; Pred. No. 1.4e-07;
Matches 39; Conservative 0; Mismatches 5; Indels 13; Gaps 4;
QY 1 AKKYAKK--AKAEKA-----KKAYKAAEAKK--AAKYKAAAEKAAAEKAAEAYEA 45
Db 1 AKKYAKKAYAKKAAKAAKAAKAYKAAEAKKAAEAKY--KAAEAKAAAEKAAEAYEA 56

RESULT 4
US-09-816-989A-5
; Sequence 5, Application US/09816989A
; Patent No. US20020115103A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816, 989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-5
Query Match 61.5%; Score 131; DB 9; Length 77;
Best Local Similarity 50.6%; Pred. No. 4.8e-07;
Matches 39; Conservative 0; Mismatches 6; Indels 32; Gaps 3;
QY 1 AKKYAKK--AKAEKA-----KKAYKAAEAKKAAKYE----- 29
Db 1 AKKYAKKAYAKKAAKAAKAAKAAKAAKAYKAAEAKKAAEAKKAYKAAKAAEAKKAYKAAEAKKAYKAAEAK 60
QY 30 -KAAAEKAAAEKAAEAYEA 45
Db 61 YKAEAAKAAAEKAAEAYEA 77

RESULT 5
US-09-816-989A-6
; Sequence 6, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816, 989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 6
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-6
Query Match 59.4%; Score 126.5; DB 9; Length 86;
Best Local Similarity 45.3%; Pred. No. 1.6e-06;
Matches 39; Conservative 0; Mismatches 6; Indels 41; Gaps 3;
QY 1 AKKYAKK--AKAEKA-----KKAYKAAEAKKAAKYE----- 29
Db 1 AKKYAKKAYAKKAAKAAKAAKAAKAAKAYKAAEAKKAAEAKKAAEAKKAYKAAEAKKAYKAAEAKKAYKAAEAK 60
QY 30 -----KAAAEKAAAEKAAEAYEA 45
Db 61 YKAEAAKAYKAAEAKAAAEKAAEAYEA 86
```

```

; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55748
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55748

Query Match          50.9%; Score 108.5; DB 15; Length 428;
Best Local Similarity 60.4%; Pred. No. 0.00079;
Matches 32; Conservative 2; Mismatches 8; Indels 11; Gaps 2;

QY      1 AKKYA-----KKAAAEKAKYKAAE-----AKKAAYEKAAAEKAAAEAA 42
Db      199 AKAAADAQKAAEAAKAAQAAEAKAAAEAAKAAAEKAAAEKAAAEKAAAE 251

RESULT 8
US-09-816-989A-1
; Sequence 1, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WE
; TITLE OF INVENTION: AND FOR THERAPEUTIC USE
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-1

Query Match          50.7%; Score 108; DB 9; Length 35;
Best Local Similarity 64.4%; Pred. No. 6.1e-05;
Matches 29; Conservative 1; Mismatches 5; Indels 10; Gaps 2;

QY      1 AKKYAKKAAEKAKYKAAEAKKAAEKAAEKAAEKAAEKAAEKAAEAAEAA 45
Db      1 AKKYAKKEKA--AKKAYK-----KEAKAKAAEAAAEKAAEAAEAA 35

RESULT 9
US-10-282-122A-56483
; Sequence 56483, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.

```

```

RESULT 11
US-10-282-122A-60543
; Sequence 60543, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

```

## RESULT 11

RESOL 11  
IIS-10-282-122A-60543

2896	Local Similarity	93.5%	Acc.: NO; 0.011,
Matches	32;	Conservative	2; Mismatches 9; Indels 6; Gaps 3;

Db 247 KKA AEKAAAEKAAADKKAAAEKAAADKKAAA-AKAAAEKAAAKAAAEAA 294

APPLICANT: XU, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; CURRENT APPLICATION NUMBER: US/10/282,122A

; COARENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR APPLICATION NUMBER: 60/206,848

;; PRIOR FILING DATE: 2000-05-26

; PRIOR FILING DATE: 2000-09-06  
; PRIOR AFFILIATION NUMBER: 00/250,333

; PRIOR FILING DATE: 2000-09-09

;; PRIOR FILING DATE: 2000-10-23  
;; PRIOR APPLICATION NUMBER: 60/253,625

;; PRIOR FILING DATE: 2000-11-27  
;; PRIOR APPLICATION NUMBER: 60/257,931

;; PRIOR FILING DATE: 2000-12-22  
: PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed -  
 : NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 60543
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60543

Query Match          44.4%; Score 94.5; DB 15; Length 239;
Best Local Similarity 56.2%; Pred. No. 0.014;
Matches 27; Conservative 7; Mismatches 9; Indels 5; Gaps 2;

QY 2 KVAKAKAEKAKK---AVKAAEAKK--AAKYEKAAAEKAAAEKAAAE 44
DB 124 KAAAEKAEADKKQBEDAVKAAANAKKEQAEAEKAAADKAAAEKAAAE 171

RESULT 12
US-10-051-643-201
; Sequence 201, Application US/10051643
; Publication No. US20020197265A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
; TITLE OF INVENTION: System using Mycobacterium Vaccae
; FILE REFERENCE: 11000.1008c2
; CURRENT APPLICATION NUMBER: US/10/051,643
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US09/156,181
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 08/996,624
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 201
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-10-051-643-201

Query Match          42.7%; Score 91; DB 13; Length 223;
Best Local Similarity 59.6%; Pred. No. 0.03;
Matches 28; Conservative 5; Mismatches 12; Indels 2; Gaps 2;

QY 1 AKKYA-KKAKAEKAKKAYKAAAEKAAKAEKAAKAEKAAAEKAAAE 45
DB 137 AKKAATKAPAKKATAAKKAAKATAAKKAAKAPAKKAPAKKAAATKA 183

RESULT 13
US-10-205-979-52
; Sequence 52, Application US/10205979
; Publication No. US20030147861A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; APPLICANT: Abernethy, Nevin
; TITLE OF INVENTION: Compounds and Methods for the Modulation
; TITLE OF INVENTION: of Immune Responses
; FILE REFERENCE: 11000.10630
; CURRENT APPLICATION NUMBER: US/10/205,979
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/308,446
; PRIOR FILING DATE: 2001-07-26
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-10-205-979-52

Query Match          42.7%; Score 91; DB 14; Length 223;
```

```
Best Local Similarity 59.6%; Pred. No. 0.03;
Matches 28; Conservative 5; Mismatches 12; Indels 2; Gaps 2;

QY 1 AKKYA-KKAKAEKAKKAYKAAAEKAAKAEKAAKAEKAAAEKAAAE 45
DB 137 AKKAATKAPAKKATAAKKAAKATAAKKAAKAPAKKAPAKKAAATKA 183

RESULT 14
US-10-229-567-27
; Sequence 27, Application US/10229567
; Publication No. US20030092080A1
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; APPLICANT: Cohavy, Offer
; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of
; TITLE OF INVENTION: Ulcerative Colitis, and Clinical Subtypes Thereof, Using
; TITLE OF INVENTION: Microbial UC PANCA antigens
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/10/229,567
; APPLICATION NUMBER: US 09/041,889
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/417,264
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 09/041,889
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-PM 3006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-10-229-567-27

Query Match          42.0%; Score 89.5; DB 14; Length 214;
Best Local Similarity 55.6%; Pred. No. 0.042;
Matches 25; Conservative 4; Mismatches 13; Indels 3; Gaps 1;

QY 1 AKKYAKKAEKAKKAYKAAAEKAAKAEKAAKAEKAAAEKAAAE 45
DB 111 AKKVAKKAPAKKATAAKKAAKATAKAPA---RKAATKAPAKKAAATKA 152

RESULT 15
US-10-282-122A-62547
; Sequence 62547, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
```



Result No.	Score	Query Match	Length	DB	ID	Description
1	213	100.0	45	4	US-09-405-743A-2	Sequence 2, Appli
2	138	64.8	109	4	US-09-405-743A-7	Sequence 7, Appli
3	134.5	63.1	56	4	US-09-405-743A-3	Sequence 3, Appli
4	131	61.5	77	4	US-08-405-743A-5	Sequence 5, Appli
5	126.5	59.4	86	4	US-09-405-743A-6	Sequence 6, Appli
6	120.5	56.6	66	4	US-09-405-743A-4	Sequence 4, Appli
7	108	50.7	35	4	US-09-405-743A-1	Sequence 1, Appli
8	95	44.6	469	4	US-09-489-039A-13565	Sequence 13565, A
9	91	42.7	223	3	US-09-095-855-201	Sequence 201, App
10	91	42.7	223	4	US-09-205-426-201	Sequence 201, App
11	89.5	42.0	214	3	US-09-041-889-27	Sequence 27, Appl
12	89.5	42.0	214	4	US-09-417-264-27	Sequence 27, Appl
13	86	40.4	148	4	US-09-248-796A-26989	Sequence 26989, A
14	86	40.4	407	4	US-09-252-991A-29581	Sequence 29581, A
15	85	39.9	176	4	US-09-248-796A-18922	Sequence 18922, A
16	81.5	38.3	497	4	US-09-134-000C-5990	Sequence 5990, Ap
17	78	36.6	316	4	US-09-252-991A-32957	Sequence 32957, A
18	77.5	36.6	700	4	US-09-107-532A-5094	Sequence 5094, Ap
19	77	36.2	103	3	US-09-041-889-39	Sequence 39, Appl
20	77	36.2	103	4	US-09-417-264-39	Sequence 39, Appl
21	77	36.2	116	3	US-09-041-889-38	Sequence 38, Appl
22	77	36.2	116	4	US-09-417-264-38	Sequence 38, Appl
23	77	36.2	158	3	US-09-041-889-40	Sequence 40, Appl
24	77	36.2	158	4	US-09-417-264-40	Sequence 40, Appl
25	77	36.2	222	3	US-09-041-889-3	Sequence 3, Appli
26	77	36.2	222	3	US-08-837-058-3	Sequence 3, Appli
27	77	36.2	222	4	US-09-417-264-3	Sequence 3, Appli







**D<sup>b</sup>** 137 AKKAATKAAPAKKATAAKKAAPAKKATAAKKAAPAKKATAAKKAATKA 183

## RESULT 11

US-09-041-889-27  
; Sequence 27, Application US/09041889  
; Patent No. 6033864  
; GENERAL INFORMATION:  
; APPLICANT: Braun, Jonathan  
; APPLICANT: Cohavy, Ofer  
; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of  
; TITLE OF INVENTION: Ulcerative Colitis, and Clinical Subtypes Thereof, Using  
; TITLE OF INVENTION: Microbial UC pANCA antigens  
; NUMBER OF SEQUENCES: 41  
; CORRESPONDENCE ADDRESS:

ADDRESS: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA

ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/041,889  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/837,058  
FILING DATE: 11-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-PM 3006  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 214 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-041-889-27

Query Match	42.0%	Score 89.5;	DB 3;	Length 214;
Best Local Similarity	55.6%	Pred. No. 0.013;		
Matches 25;	Conservative	4;	Mismatches 13;	Indels 3;
Gaps 1;				

Qy	1	AKYAKKAAEAKKAYKAAEAKKAAKYEKAAAEKAAAEKAAEAYEA	45
		:         :           :   :	
Db	111	AKKAKKAPAKKATYKAAKKAATKAPA---RKAAATKAPAKKAATKA	151

## RESULT 12

US-09-417-264-27  
 ; Sequence 27, Application US/09417264  
 ; Patent No. 6537768  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Braun, Jonathan  
 ; APPLICANT: Cohavy, Offer  
 ; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of  
 ; TITLE OF INVENTION: Ulcerative Colitis, and Clinical Subtypes Thereof, Using  
 ; TITLE OF INVENTION: Microbial UC pANCA antigens  
 ; NUMBER OF SEQUENCES: 41  
 ; CORRESPONDENCE ADDRESS:

ADDRESS: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California

```

; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/417,264
; FILING DATE:
;

```

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/041,889  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-PM 3006  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 214 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-417-264-27

Query Match	42.0%;	Score 89.5;	DB 4;	Length 214;
Best Local Similarity	55.6%;	Pred. No. 0.013;		
Matches 25;	Conservative	4;	Mismatches 13;	Indels 3;
				Gaps 1;

Qy 1 AKTVAKKAEEAKKAYKAAEAKKAAYEKAAAEKAAAEKAAEYEA 45

Db 111 AKTVAKKAPAKKATKAAKKAATKAPAAA---RKAATKAPAKKAATKA 152

## RESULT 13

```

US-09-248-796A-26989
; Sequence 26989, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RE
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 26989
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-26989

```

Query Match	40.4%	Score 86;	DB 4;	Length 148;
Best Local Similarity	48.9%;	Pred. No. 0.021;		
Matches 23;	Conservative	10;	Mismatches 12;	Indels 2;
Gaps 1;				

[illegible]

RESULT 14  
US-09-252-991A-29581  
; Sequence 29581, Application US/09252991A  
; Patent No. 6551795

```
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29581
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29581

Query Match 40.4%; Score 86; DB 4; Length 407;
Best Local Similarity 51.1%; Pred. No. 0.06;
Matches 24; Conservative 6; Mismatches 15; Indels 2; Gaps 1;

QY 1 AKKYAKKAKAEKAKKA--YKAAEAKKAKYKAAAEKAAKAAKAAEAA 45
DB 185 AQAARAKKAEAKKAAEAKAAEQKQADIAKRAEDEAKKAAEDA 231

RESULT 15
US-09-248-796A-18922
; Sequence 18922, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18922
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18922

Query Match 39.9%; Score 85; DB 4; Length 176;
Best Local Similarity 50.0%; Pred. No. 0.033;
Matches 27; Conservative 7; Mismatches 8; Indels 12; Gaps 3;

QY 2 KKYAKKAKAEKAKKAY-----KAAE---AKKAAKYEKAAAEKAAKAAEAA 45
DB 81 KLFKEKAKKEKAKWEKSGSRKAAAEAAKAA--EEAAAKKAAEEAAALKA 132
```

Search completed: December 14, 2004, 05:50:16  
Job time : 13.7694 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 12.8776 Seconds  
(without alignments)  
418.411 Million cell updates/sec

Title: US-10-792-311-3  
Perfect score: 266  
Sequence: 1 AKYAKKEKAYAKAEKAAK.....EAKYKAEAKAAKEAYEA 56  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues 283416  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: Piri:\*  
2: Piri:\*  
3: Piri:\*  
4: Piri:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	117	44.0	421	JV0057	tola protein - Esc
2	116.5	43.8	239	AE1317	hypothetical prote
3	113	42.5	394	F90725	membrane spanning
4	113	42.5	394	G85576	membrane spanning
5	110.5	41.5	376	AG0592	tola protein (impo
6	110	41.4	1701	T09127	probable erythrocy
7	105.5	39.7	372	G64064	outer membrane int
8	102.5	38.5	210	A25550	histone H1 - sea u
9	102.5	38.5	347	E83525	Tola protein PA097
10	102	38.3	211	A28100	histone H1-beta, e
11	100.5	37.8	1390	S51364	sperm tail-specifi
12	99.5	37.4	206	HSR1R	histone H1 - rainb
13	98.5	37.0	243	AE1689	hypothetical prote
14	98.5	37.0	311	T17698	hypothetical prote
15	97.5	36.7	206	S09388	histone H1 - sea u
16	97	36.5	219	E80110	repetitive protein
17	96	36.1	384	B43592	outer membrane pro
18	95.5	35.9	217	A26721	histone H1-gamma,
19	95.5	35.9	248	HSUR1P	histone H1, gonada
20	95.5	35.9	1403	T11583	probable translati
21	95	35.7	388	AC0138	Tola colicin impor
22	94.5	35.5	214	G70673	probable hupB - My
23	94.5	35.5	328	A44993	cytosolic repetiti
24	94	35.3	433	S25194	zuotin - yeast (Sa
25	94	35.3	703	T48600	kinase-like protei
26	92.5	34.8	182	S61926	histone H1 homolog
27	92.5	34.8	1128	T30296	R27-2 protein - Tr
28	92	34.6	220	A28456	histone H1.10 - ch
29	92	34.6	236	S22322	histone H1 - wheat

ALIGNMENTS

RESULT 1

JV0057  
tola protein - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C>Date: 07-Sep-1990 #sequence revision 07-Sep-1990 #text\_change 09-Jul-2004  
C:Accession: JV0057; B64810  
R:Levengood, S.K.; Webster, R.E.  
J. Bacteriol. 171, 6600-6609, 1989  
A:Title: Nucleotide sequences of the tola and toIB genes and localization of their products  
A:Reference number: JV0057; MUID:90078104; PMID:2687247  
A:Accession: JV0057  
A:Molecule type: DNA  
A:Residues: 1-421 <LEV>  
A:Cross-references: UNIPROT:PI9934; GB:M28232; NID:G148018; PIDN:AAA24683.1; PID:G148019  
A:Experimental source: strain JM105  
A:Note: The authors translated the initiation codon GTG for residue 1 as Val  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: B64810  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-421 <BLAT>  
A:Cross-references: GB:AE000177; GB:U00096; NID:G1786955; PIDN:AAC73833.1; PID:G1786960;  
A:Experimental source: strain K-12, substrain MG1655  
C:Comment: tola and toIB proteins are necessary for colicing E2, E3, A, and K to reach the  
C:Genetics:  
A:Gene: tola  
A:Map position: 17 min  
A:Start codon: GTG  
C:Keywords: nucleotide binding; P-loop; transmembrane protein  
F:14-34/Domain: transmembrane #status predicted <MSS>  
F:78-30/Domain: helical #status predicted <HSR>  
F:355-362/Region: nucleotide-binding motif A (P-loop)

Query Match 44.0%; Score 117; DB 2; Length 421;  
Best Local Similarity 54.7%; Pred. No. 0.0048;  
Matches 35; Conservative 5; Mismatches 10; Indels 14; Gaps 2;

QY 7 KEKAYAKAEKAAKAEKAYK-----AAEAKKKAE-----KYKAEAKAAKAE 52

Db 148 KAEADAKAEAEAKKAAADAKKKAEAEAKAAAEAKKAEAAALKKKAEAEAEAAEA 207

QY 53 AYEAE 56

Db 208 RKKA 211

RESULT 2

AE1317



[illegible]

C;Species: Drosophila hydei  
C;Date: 19-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: S51364; S34154  
R;Neeseen, J.; Padmanabhan, S.; Buenemann, H.  
Eur. J. Biochem. 225, 1089-1095, 1994  
A;Title: Tandemly arranged repeats of a novel highly charged 16-amino-acid motif represent alpha-helical rods within the extremely elongated spermatozoa of Drosophila hydei.  
A;Reference number: S51364; MUID:95045538; PMID:7957199  
A;Accession: S51364  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 1-1390 <NEE>  
A;Cross-references: UNIPROT:Q08696; EMBL:X73481  
R;Neeseen, J.; Heinlein, U.A.O.; Buenemann, H.  
submitted to the EMBL Data Library, June 1993  
A;Reference number: S34153  
A;Accession: S34154  
A;Molecule type: DNA  
A;Residues: 1-163, 'E', 164-236, 'Q', 237-254, 257-320, 'E', 321-1390 <NEW>  
A;Cross-references: EMBL:X73481; NID:g313201; PID:g313202  
C;Genetics:  
A;Gene: mat101(2)  
A;Cross-references: FlyBase:FBgn0011816

Query Match 37.8%; Score 100.5; DB 2; Length 1390;  
Best Local Similarity 49.2%; Pred. No. 0.29;  
Matches 29; Conservative 5; Mismatches 18; Indels 7; Gaps 2;

QY 2 KKYAKKEK-AYAKKAEKAAK-----AEAKYKAAEAKKAAKAAKAAKAAKAA 53  
Db 511 KKLAKKEKETAEKKCKEAKAKRKEAEKKCKEAKAAKRAAEKKCKEAKKCKEAKKRAKKEA 569

RESULT 12  
HSTR1R  
histone H1 - rainbow trout  
C;Species: Oncorhynchus mykiss (rainbow trout)  
C;Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 09-Jul-2004  
C;Accession: A02584  
R;Mezquita, J.; Connor, W.; Winkfein, R.J.; Dixon, G.H.  
J. Mol. Evol. 21, 209-219, 1985  
A;Title: An H1 histone gene from rainbow trout (Salmo gairdnerii).  
A;Reference number: A02584; MUID:95264847; PMID:6443128  
A;Accession: A02584  
A;Molecule type: DNA  
A;Residues: 1-206 <MEZ>  
A;Cross-references: UNIPROT:P06350; GB:X02624; NID:g64323; PIDN:CAB37646.1; PID:g4468016  
C;Superfamily: histone H1  
C;Keywords: chromosomal protein; DNA binding; nucleosome  
F;1-44/Region: flexible nose  
F;45-117/Region: globular head  
F;118-206/Region: flexible tail

Query Match 37.4%; Score 99.5; DB 1; Length 206;  
Best Local Similarity 46.7%; Pred. No. 0.08;  
Matches 28; Conservative 7; Mismatches 18; Indels 7; Gaps 1;

QY 1 AKKYAKKEKAYAKAEK-----AKKAEKAYKAAEAKKAAKAAKAAKAAKAA 53  
Db 117 AKKVAKKPAAPKPKVAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 176

RESULT 13  
AE1689  
hypothetical protein homolog lin2055 [imported] - Listeria innocua (strain Clip11262)  
C;Species: Listeria innocua  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AE1689  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloechel, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A;Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AE1689  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-243 <GLA>  
A;Cross-references: UNIPROT:O92A67; GB:AL592022; PIDN:CAC97285.1; PID:g16414556; GSPDB:G  
A;Experimental source: strain Clip11262  
C;Genetics:  
A;Gene: lin2055

Query Match 37.0%; Score 98.5; DB 2; Length 243;  
Best Local Similarity 50.0%; Pred. No. 0.11;  
Matches 28; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

QY 1 AKKYAKKEKAYAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 55  
Db 116 AEKAAEKKAAEKAADKKSODEAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 171

RESULT 14  
TI17698  
hypothetical protein A208R - Chlorella virus PBCV-1  
C;Species: Chlorella virus PBCV-1  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: TI17698  
R;Graves, M.V.; Van Etten, J.L.  
submitted to the EMBL Data Library, May 1999  
A;Reference number: Z18806  
A;Accession: TI17698  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-311 <GRA>  
A;Cross-references: UNIPROT:O84528; EMBL:U42580; NID:g4028896; PIDN:AAC96576.1  
A;Experimental source: specific host Chlorella strain NC64A  
C;Genetics:  
A;Note: A208R

Query Match 37.0%; Score 98.5; DB 2; Length 311;  
Best Local Similarity 43.9%; Pred. No. 0.13;  
Matches 29; Conservative 8; Mismatches 18; Indels 11; Gaps 2;

QY 2 KKYAKKEKAYAK-----KAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 50  
Db 72 KERVAKAAAREAEKERVAKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEK 131

QY 51 EAAYEA 56  
Db 132 EAAKAA 137

RESULT 15  
S09388  
histone H1 - sea urchin (Parechinus angulosus)  
C;Species: Parechinus angulosus (angulate urchin)  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S09388  
R;Hill, C.S.; Martin, S.R.; Thomas, J.O.  
EMBO J. 8, 2591-2599, 1989  
A;Title: A stable alpha-helical element in the carboxy-terminal domain of free and chroma  
A;Reference number: S09388; MUID:90060019; PMID:2583125  
A;Accession: S09388  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-206 <HIL>  
A;Cross-references: UNIPROT:Q7M409  
C;Superfamily: histone H1  
C;Keywords: chromosomal protein

Query Match 36.7%; Score 97.5; DB 2; Length 206;  
Best Local Similarity 52.6%; Pred. No. 0.12;  
Matches 30; Conservative 3; Mismatches 23; Indels 1; Gaps 1;



Qy 1 AKKYAKKEKAYAK-KAEKAAKKAEAKYKAAEAKKKAEAKYKAEAKKAAAEAKKAEAYEA 56  
|: ||| ||| ||| ||| ||| ||| : ||| ||| :  
Db 132 ABAAKKAAKAAKRAALAKKAAKRAATKAKKAKKKKTAAKKAKKAAKKA 188

Search completed: December 14, 2004, 05:52:13  
Job time : 13.8776 secs

***This Page Blank (uspto)***









RP	SEQUENCE FROM N.A.
RC	STRAIN=17XNL;
RX	PubMed=12368865;
RA	Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perteau M.,
RA	Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA	Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA	Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA	Cho J.K., Quackenbush J., Sedegah M., Shoabli A., Cummings L.M.,
RA	Florens L., Yates F.R. III, Raine J.D., Sindén R.E., Harris M.A.,
RA	Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA	van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA	Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA	Carucci D.J.;
RT	"Genome sequence and comparative analysis of the model rodent malaria
RT	parasite Plasmodium yoelii yoelii";
RL	Nature 419:512-519(2002).
CC	-!- CAUTION: The sequence shown here is derived from an
CC	EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC	preliminary data.
DR	EMBL; AABL01001968; EAA18109.1; -.
DR	GO; GO:0016021; C:integral to membrane; IEA.
DR	GO; GO:0004872; F:receptor activity; IEA.
DR	GO; GO:0009405; P:pathogenesis; IEA.
DR	InterPro; IPR008602; Duffy_binding.
DR	Pfam; PF05424; Duffy_Binding; 1.
SQ	SEQUENCE 1701 AA; 197910 MW; 03BE665BAE45C669 CRC64;
Query Match	43.4%; Score 115.5; DB 2; Length 1701;
Best Local Similarity	59.3%; Pred. No. 0.16; Indels 9; Gaps 3;
Matches	35; Conservative 4; Mismatches 11; Indels 9; Gaps 3;
Qy	1 AKKYAKKKA-YAKYAKKKA-KAYAKYAKKKA-A-----KYKAEAAKAAK 51
Db	1270 AKKAEKKKAAKAAEKKEKKAAEA-AKKAEEKKKAAKAAEKKEKKAAEKKE 1327
RESULT 10	
ID	Q7C2Q4 PRELIMINARY; PRT; 413 AA.
AC	O7C2Q4;
DT	05-JUL-2004 (TrEMBLrel. 27, Created)
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE	Membrane spanning protein.
DN	Name=tOla; OrderedLocusNames=S0571;
OS	Shigella flexneri.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC	Enterobacteriaceae; Shigella.
OX	NCBI_TaxID=623;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=2457T;
RX	MEDLINE=22590274; PubMed=12704152;
RA	Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA	Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA	Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA	Schwartz D.C., Blattner F.R.;
RT	"Complete genome sequence and comparative genomics of Shigella
RT	flexneri serotype 2a strain 2457T.";
RL	Infect. Immun. 71:2775-2786(2003).
DR	EMBL; AE016979; AAP16075.1; -.
DR	InterPro; IPR010528; ToLA.
DR	Pfam; PF06519; ToLA; 1.
SQ	SEQUENCE 413 AA; 42355 MW; 93E10F2C5DE60DE8 CRC64;
Query Match	43.2%; Score 115; DB 2; Length 413;
Best Local Similarity	51.4%; Pred. No. 0.055;
Matches	36; Conservative 5; Mismatches 15; Indels 14; Gaps 2;
Qy	1 AKKYAKKKA-YAKYAKKKA-KAYAKYAKKKA-A-----AAEAKKKAAE-----KYKAEAAK 46
Db	134 AEVAAAACAADAACAEEAAKAAADAKKCAEAAKAAEAAKAAEAAALKKCAEAAE 193
RESULT 12	
ID	Q8FUJ1 PRELIMINARY; PRT; 421 AA.
AC	Q8FUJ1;
DT	01-WAR-2003 (TrEMBLrel. 23, Created)
DT	01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DE	Membrane spanning protein.
DN	Name=tOla; OrderedLocusNames=c0818;
OS	Escherichia coli O6.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC	Enterobacteriaceae; Escherichia.
OX	NCBI_TaxID=217992;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=06:H1 / CFT073 / ATCC 700928;
RX	MEDLINE=22388234; PubMed=12471157;
RA	Weich R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA	Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA	Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA	Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT	"Extensive mosaic structure revealed by the complete genome sequence
RT	of uropathogenic Escherichia coli.";

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016757; AAN79291.1; -.
DR HSSP; P19934; ITOL.
DR InterPro; IPR010528; Tola.
DR Pfam; PF06519; Tola; 1.
KW Complete proteome.
SQ SEQUENCE 421 AA; 43184 MW; DB296626F056D385 CRC64;

Query Match 43.2%; Score 115; DB 2; Length 421;
Best Local Similarity 56.2%; Pred. No. 0.056;
Matches 36; Conservative 6; Mismatches 8; Indels 14; Gaps 3;

QY 7 KEKAYAKAEKAA-----KKAEEKAYKAA-EAKKKAEEA-----KYKAAEAAKAAKAA 52
Db 148 KAEADAKAAEAAKKAADAAKKAEEAAKAAEAAVQAQKAAEAAALKKKAAEAAEAAEA 207
QY 53 AYE 56
Db 208 RKKA 211

RESULT 13
Q7W477 PRELIMINARY; PRT; 373 AA.
AC Q7W477
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Proline-rich inner membrane protein.
GN OrderedLocuNames=BPP3791;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1237;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomason N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Baeson N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40 (2003).
DR EMBL; BX640434; CAE39074.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. .; IEA.
DR GO; GO:0008565; F:protein transport activity; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR InterPro; IPR010528; Tola.
DR Pfam; PF06519; Tola; 1.
DR TIGRFAMs; TIGR01352; tonB_Cterm; 1.
KW Complete proteome.
SQ SEQUENCE 373 AA; 40092 MW; A364894DAE46E328 CRC64;

Query Match 43.0%; Score 114.5; DB 2; Length 373;
Best Local Similarity 55.4%; Pred. No. 0.055;
Matches 31; Conservative 8; Mismatches 16; Indels 1; Gaps 1;

QY 1 AKKYAKKEKAYAKA-EKAAKAEKAYKAAEAKKAEAKYKAAEAAKAAKAAEAAE 55
Db 182 AEKAAEKQAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 237

RESULT 14
Q937K4 PRELIMINARY; PRT; 395 AA.
AC Q937K4
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tola protein.
GN Name=tola;
OS Erwinia chrysanthemi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=556;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=3937;
RA Ray M.C., Vianney A., Cotte-pattat N., Lazzaroni J.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ297885; CAC82708.1; -.
DR HSSP; P19934; ITOL.
DR InterPro; IPR010528; Tola.
DR Pfam; PF06519; Tola; 1.
SQ SEQUENCE 395 AA; 41601 MW; 3C0C1DC12E181013 CRC64;

Query Match 43.0%; Score 114.5; DB 2; Length 395;
Best Local Similarity 54.2%; Pred. No. 0.058;
Matches 32; Conservative 11; Mismatches 13; Indels 3; Gaps 2;

QY 1 AKKYAKKE-KAYAKAEKAAKAA-EAKYKAAEAKKAEAKYKAAEAAKAAEAAE 56
Db 182 AKKAAEEAAKTKAAAAAEAKKAAEEAAKAAADAAKQKAAEEAAKAAEAAEAA 240

RESULT 15
Q8ZQT6 PRELIMINARY; PRT; 407 AA.
AC Q8ZQT6
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tol protein, membrane spanning protein.
GN Name=tola; OrderedLocuNames=STM0747;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856 (2001).
DR EMBL; AE008730; AAL19691.1; -.
DR HSSP; P19934; ITOL.
DR InterPro; IPR010528; Tola.
DR Pfam; PF06519; Tola; 1.
KW Complete proteome.
SQ SEQUENCE 407 AA; 41865 MW; 2534352116602D75 CRC64;

Query Match 42.7%; Score 113.5; DB 2; Length 407;
Best Local Similarity 49.3%; Pred. No. 0.072;
Matches 37; Conservative 10; Mismatches 9; Indels 19; Gaps 4;

QY 1 AKKYAKKEKAYAKAEKAA-----KKAEEKAYK-AAEAKKAAE-----KYKAAE 44
Db 123 AKLAQQQQQAEAAKAAADAAKKAEEAAKAAADAAKKAEEAAEAAVKAADAAKKAEEA 182
QY 45 AKAAA---KEAAE 56

```



Db           |||||   |:|   ||  
          183 AKAAADAKKKAAEA 197

Search completed: December 14, 2004, 06:10:16  
Job time : 63.8523 secs

***This Page Blank (uspto)***











***This Page Blank (uspto)***







[illegible]

```
SQ SEQUENCE 413 AA; 42355 MW; 93E10F2C5DB60DE8 CRC64;
Query Match 40.7%; Score 127.5; DB 2; Length 413;
Best Local Similarity 56.7%; Pred. No. 0.013;
Matches 38; Conservative 6; Mismatches 20; Indels 3; Gaps 2;
QY 1 AKKYAKKEKAYAKAKAEAKKAAKAEAKKAAKAEKKAEEKK-YYAAAEAKKYKAEAAKAAA 59
Db 139 AKAADAKAAEEAAKKAADAKKKAEAEAK--AAAEAKKAEVAAALKKAEAEAAA 196
QY 60 KEAAYEA 66
Db 197 AEARKKA 203
RESULT 7
Q9CM70 PRELIMINARY; PRT; 389 AA.
ID Q9CM70
AC Q9CM70;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ToLA.
GN Name=toLa; OrderedLocusNames=PM0968;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida PM70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).
DR EMBL; A5006136; AAK03052.1; -.
DR HSSP; P01096; IGMJ.
DR InterPro; IPR009148; Siba.
DR InterPro; IPR010528; ToLa.
DR InterPro; IPR000533; Tropomyosin.
DR Pfam; PF06519; ToLa; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PRINTS; PR00194; TROPOMYOSIN.
KW Complete proteome.
SQ SEQUENCE 389 AA; 42197 MW; B4032F2A2FD9E94B CRC64;
Query Match 40.4%; Score 126.5; DB 2; Length 389;
Best Local Similarity 52.0%; Pred. No. 0.014;
Matches 39; Conservative 9; Mismatches 12; Indels 15; Gaps 3;
QY 1 AKKYAKKEKAYAKAK-----KAEAKA--AKKAEAKKYKAEAKAEKKEYYAAAEAKYK 51
Db 189 AKKAEAKAEAEAKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 242
QY 52 AEAKAEEAEAYEA 66
Db 243 AEKAKADAEAAQKKA 257
RESULT 8
Q8ZQT6 PRELIMINARY; PRT; 407 AA.
ID Q8ZQT6
AC Q8ZQT6
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tol protein, membrane spanning protein.
GN Name=toLa; OrderedLocusNames=STM0747;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856 (2001).
DR EMBL; AE008730; AAL19691.1; -.
DR HSSP; P19934; 1TOL.
DR InterPro; IPR010528; ToLa.
DR Pfam; PF06519; ToLa; 1.
KW Complete proteome.
SQ SEQUENCE 407 AA; 41865 MW; 2534352116602D75 CRC64;
Query Match 39.8%; Score 124.5; DB 2; Length 407;
Best Local Similarity 47.8%; Pred. No. 0.022;
Matches 43; Conservative 6; Mismatches 10; Indels 31; Gaps 4;
QY 1 AKKYAKKEKAYAKA---KAEAKAAK-----KAEAKKYKAEAKAEKKEY 43
Db 189 AKKCAEAEAAKAAAEAKKAEAEAAKAAAEAKKADAEAAKAAAEAKKADAAAAAK---- 244
QY 44 AAEEAKYKAEAA-----KAAAEKAA 63
Db 245 AAEEAKKADAAAKAAADAKKKAAAEKAA 274
RESULT 9
HL_PARAN STANDARD; PRT; 248 AA.
ID HL_PARAN
AC P02256;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Histone H1, gonadal.
OS Parechinus angulosus (Angulate sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoidae; Echinidae; Parechinus.
OX NCBI_TaxID=7658;
RN [1]
RP SEQUENCE OF 1-84.
RX MEDLINE=80156831; PubMed=6767609;
RA Strickland W.N., Strickland M., de Groot P.C., von Holt C.,
Wittmann-Liebold B.;
RT "The primary structure of histone H1 from sperm of the sea urchin
Parechinus angulosus. 1. Chemical and enzymatic fragmentation of the
protein and the sequence of amino acids in the four N-terminal
cyanogen bromide peptides.";
RN [2]
RP SEQUENCE OF 80-248.
RX MEDLINE=80156832; PubMed=7363905;
RA Strickland W.N., Strickland M., Brandt W.F., von Holt C., Lehmann A.,
Wittmann-Liebold B.;
RT "The primary structure of histone H1 from sperm of the sea urchin
Parechinus angulosus. 2. Sequence of the C-terminal CNBr peptide and
the entire primary structure.";
RL Eur. J. Biochem. 104:567-578 (1980).
CC -1- FUNCTION: Histones H1 are necessary for the condensation of
nucleosome chains into higher order structures.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Sperm.
CC -1- SIMILARITY: Belongs to the histone H1/H5 family.
DR PIR; A91090; HSUR1P.
DR HSSP; P02259; 1HST.
DR InterPro; IPR005818; Histone_H1/H5.
DR InterPro; IPR005819; Histone_H5.
DR InterPro; IPR003216; Linkerhist_N.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
```

[illegible]

```

Db 148 AKKAAEQQADYIAKKAAEDAKKABEEAK---KAAAEBAKKAEDAKKAAEEAKKA 204
QY 59 AKEAAVEA 66
Db 205 AEDAKKA 212

RESULT 12
Q828C1 PRELIMINARY; PRT; 376 AA.
AC Q828C1; Q7C8P3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Tola protein.
DE Tola protein.
GN Name=tola; OrderedLocusNames=STV0793, t2129;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RA "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TV2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodyonanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18."
RL J. Bacteriol. 185:2330-2337(2003).
DR EMBL; AL627268; CAD05209.1; -.
DR EMBL; AE016841; AAO69743.1; -.
DR HSSP; P19934; ITOI.
DR InterPro; IPR010528; Tola.
DR Pfam; PF06519; Tola; 1.
KW Complete proteome.
SQ SEQUENCE 376 AA; 38804 MW; EC21P2C4767A8A42 CRC64;

Query Match 39.1%; Score 122.5; DB 2; Length 376;
Best Local Similarity 51.3%; Pred. No. 0.029;
Matches 40; Conservative 4; Mismatches 13; Indels 21; Gaps 3;

QY 1 AKKYAKKEKAYAKA---KKAEEAAK-----KAKAAKKYAKAAKAEKKEY 43
Db 173 AKKAAEAAKAAAEAKKAAEAAKAAADAKKKADAAEAAKAAAEAKKADAAAK----- 228

QY 44 AAAEAKYKAAEAAKAAAE 61
Db 229 AAADAKKAAAEKAAAE 246

RESULT 13
Q6CDX0 PRELIMINARY; PRT; 713 AA.
AC Q6CDX0;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Similar to sp|P18678 Neurospora crassa Glucan synthase-1.

```

```

GN ORFNames=YAL10B20570g;
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG GENOLEVURES;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanloter F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts."
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; CR382128; CAG83395.1; -.
SQ SEQUENCE 713 AA; 77807 MW; 5DD84D2C34CF55AC CRC64;

Query Match 39.0%; Score 122; DB 2; Length 713;
Best Local Similarity 50.7%; Pred. No. 0.056;
Matches 35; Conservative 8; Mismatches 18; Indels 8; Gaps 2;

QY 2 KKVAKEKAYAKAKKAE-----AKAAKKAAKAEK--KYAKAAKAEKVEAAAEAKVKA 53
Db 571 EKEEKAATAAKAAKAEELKKEVAKAAKAEKQAATAAAEAKAAKAAEADASKKVE 630

QY 54 AAKAAAKEA 62
Db 631 AEKAAAEES 639

RESULT 14
Q74W64 PRELIMINARY; PRT; 401 AA.
AC Q74W64;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Tola colicin import membrane protein.
GN Name=tola; OrderedLocusNames=YPI033;
OS Versinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017130; AAS61283.1; -.
DR InterPro; IPR010528; Tola.
DR Pfam; PF06519; Tola; 1.
SQ SEQUENCE 401 AA; 41868 MW; 43D682DE91CF1301 CRC64;

Query Match 38.8%; Score 121.5; DB 2; Length 401;

```



***This Page Blank (uspto)***









```

RESULT 7
US-09-248-796A-26989
; Sequence 26989, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 26989
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-26989

Query Match          41.5%; Score 110.5; DB 4; Length 148;
Best Local Similarity 57.7%; Pred. No. 0.00028;
Matches 30; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

QY      5 AKKEKAYAKKAEEAAKKAEAKYKAAEAKKKAEAKYKAAEAKKAAEAAYEA 56
       ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      6 AKKAAEAKKKEBEAKAE-EAKKAEAKKVEEAAKKAEAKKAEFEARKKA 56
       ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

RESULT 8
US-09-248-796A-26122
; Sequence 26122, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 26122
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (42)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-26122

Query Match          41.4%; Score 110; DB 4; Length 264;
Best Local Similarity 54.9%; Pred. No. 0.00056;
Matches 28; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY      6 KKEKAYAKKAEEAAKKAEAKYKAAEAKKKAEAKYKAAEAKKAAEAAYEA 56
       ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db     122 KKEEAKKKEEAKKAAEAKKAEAKKVEEAAKKAEAKKAEFEARKKA 172
       ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

RESULT 9
US-09-489-039A-13565
; Sequence 13565, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLBESIELLA
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 15.7131 Seconds  
(without alignments)  
236.351 Million cell updates/sec

Title: US-10-792-311-3  
Perfect score: 266  
Sequence: 1 AKKYAKKQKAYAKKAKEAKA.....EAKYKAAKAAKAAKEAAVEA 56

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/ECTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	266	100.0	56	4 US-09-405-743A-3	Sequence 3, Appli
2	245.5	92.3	77	4 US-09-405-743A-5	Sequence 5, Appli
3	233	87.6	86	4 US-09-405-743A-6	Sequence 6, Appli
4	198	74.4	66	4 US-09-405-743A-4	Sequence 4, Appli
5	180.5	67.9	109	4 US-09-405-743A-7	Sequence 7, Appli
6	134.5	50.6	45	4 US-09-405-743A-2	Sequence 2, Appli
7	110.5	41.5	148	4 US-09-248-796A-26989	Sequence 26989, A
8	110	41.4	264	4 US-09-248-796A-26122	Sequence 26122, A
9	108.5	40.8	469	4 US-09-489-039A-13565	Sequence 13565, A
10	104.5	39.3	35	4 US-09-405-743A-1	Sequence 1, Appli
11	104	39.1	56	3 US-08-993-008A-6	Sequence 6, Appli
12	104	39.1	100	2 US-08-460-890A-64	Sequence 64, Appli
13	104	39.1	100	3 US-08-167-641C-64	Sequence 64, Appli
14	104	39.1	100	3 US-08-460-971A-64	Sequence 64, Appli
15	104	39.1	100	3 US-08-462-040-64	Sequence 64, Appli
16	103	38.7	223	3 US-09-095-855-201	Sequence 201, Appli
17	103	38.7	223	4 US-09-205-426-201	Sequence 201, Appli
18	103	38.7	361	4 US-09-543-681A-5390	Sequence 5390, Appli
19	102.5	38.5	407	4 US-09-252-991A-29581	Sequence 29581, A
20	101	38.0	1507	3 US-08-929-329-5	Sequence 5, Appli
21	100	37.6	214	4 US-09-328-352-5169	Sequence 5169, Appli
22	100	37.6	700	4 US-09-107-532A-5094	Sequence 5094, Appli
23	99.5	37.4	48	3 US-08-993-008A-5	Sequence 5, Appli
24	98	36.8	176	4 US-09-248-796A-18922	Sequence 18922, A
25	97	36.5	239	4 US-09-248-796A-25676	Sequence 25676, A
26	94.5	35.5	214	3 US-09-041-889-27	Sequence 27, Appli
27	94.5	35.5	214	4 US-09-417-264-27	Sequence 27, Appli

28	94	35.3	433	1 US-08-346-849-2	Sequence 2, Appli
29	94	35.3	433	2 US-08-293-284A-2	Sequence 2, Appli
30	94	35.3	433	4 US-08-898-300-2	Sequence 2, Appli
31	94	35.3	921	4 US-09-543-681A-5734	Sequence 5734, Ap
32	92.5	34.8	643	2 US-08-216-894-8	Sequence 8, Appli
33	92.5	34.8	643	3 US-09-115-746-8	Sequence 8, Appli
34	91.5	34.4	60	1 US-08-346-849-16	Sequence 16, Appli
35	91.5	34.4	60	2 US-08-293-284A-16	Sequence 16, Appli
36	91.5	34.4	60	4 US-08-898-300-16	Sequence 16, Appli
37	91.5	34.4	472	2 US-08-216-894-10	Sequence 10, Appli
38	91.5	34.4	472	3 US-09-115-746-10	Sequence 10, Appli
39	91.5	34.4	564	2 US-08-216-894-2	Sequence 2, Appli
40	91.5	34.4	564	3 US-09-115-746-2	Sequence 2, Appli
41	91	34.2	207	4 US-09-489-039A-13743	Sequence 13743, A
42	90.5	34.0	468	4 US-09-328-352-6321	Sequence 6321, Ap
43	90.5	34.0	1507	6 5268270-2	Patent No. 5268270
44	87	32.7	1231	4 US-08-714-741-41	Sequence 41, Appli
45	86.5	32.5	171	4 US-09-270-767-39148	Sequence 39148, A

ALIGNMENTS

RESULT 1  
US-09-405-743A-3  
; Sequence 3, Application US/09405743A  
; Patent No. 6514938  
; GENERAL INFORMATION:  
; APPLICANT: Yeda Research and Development Co., Ltd.  
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS  
; FILE REFERENCE: 60807-A  
; CURRENT APPLICATION NUMBER: US/09/405,743A  
; CURRENT FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 56  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
; OTHER INFORMATION: PEPTIDE  
US-09-405-743A-3

Query Match	100.0%;	Score 266;	DB 4;	Length 56;
Best Local Similarity	100.0%;	Pred. No. 3e+20;	Mismatches 0;	Indels 0;
Matches	56;	Conservative 0;	0;	Gaps 0;
QY	1	AKKYAKKQKAYAKKAKEAKA	AKKYAKKA	AKKYAKKA
Db	1	AKKYAKKQKAYAKKAKEAKA	AKKYAKKA	AKKYAKKA

RESULT 2  
US-09-405-743A-5  
; Sequence 5, Application US/09405743A  
; Patent No. 6514938  
; GENERAL INFORMATION:  
; APPLICANT: Yeda Research and Development Co., Ltd.  
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS  
; FILE REFERENCE: 60807-A  
; CURRENT APPLICATION NUMBER: US/09/405,743A  
; CURRENT FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 77  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
; OTHER INFORMATION: PEPTIDE  
US-09-405-743A-5

```
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59321
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59321

Query Match      40.8%; Score 108.5; DB 15; Length 323;
Best Local Similarity 55.7%; Pred. No. 0.01; Indels 7; Gaps 3;
Matches 34; Conservative 7; Mismatches 13;

QY 1 AKKYAKKEKAYAKKAEKAAKAEKAYK-AAEAKKKAE- ---KYKAEAAKAAAEAAAYE 55
Db 65 AKKQA--EAAAKAAAEAKKQAEAEAAKAAAEAAQKAEAAAKKAEKQAEKQAEAKQ 122

QY 56 A 56
Db 123 A 123
```

Search completed: December 14, 2004, 06:59:41  
Job time : 139.937 secs

**This Page Blank (uspto)**





```

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 75047
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Salmonella typhimurium
; NAME/KEY: MISC FEATURE
; LOCATION: (78)..(78)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (303)..(303)
; OTHER INFORMATION: X=any amino acid
; US-10-282-122A-75047

Query Match 42.7%; Score 113.5; DB 15; Length 407;
Best Local Similarity 49.3%; Pred. No. 0.0044;
Matches 37; Conservative 10; Mismatches 9; Indels 19; Gaps 4;

QY 1 AKYAKKQYAKYAKKAAKAA-----KAAKAYK-AAEAKKKAA-----KYKAA 44
DB 123 AAKLAQQQQQEEAAKAAADAKKKAAEAAKAAADAKKKAAEAAVKAAADAKKKAAEAA 182

QY 45 AKAA--KEAAYEA 56
DB 183 AKAADAKKKAAEAA 197

RESULT 12
US-10-282-122A-72645
; Sequence 72645, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 72645
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Salmonella paratyphi A
; NAME/KEY: MISC FEATURE
; LOCATION: (78)..(78)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (303)..(303)
; OTHER INFORMATION: X=any amino acid
; US-10-282-122A-72645

Query Match 42.3%; Score 112.5; DB 15; Length 387;
Best Local Similarity 53.5%; Pred. No. 0.0052;
Matches 38; Conservative 5; Mismatches 11; Indels 17; Gaps 4;

QY 1 AKYAKKQYAKYAKKAAKAAKAYK-AAEAKKKAA-----KYKAAEAAKAA 48
DB 141 AKKAEPA--AKAADAKKKAAEAAVKAAADAKKKAAEAAKAAEAAKAA 198

QY 49 A---KEAAYEA 56
DB 199 AEAKKKAAEAA 209

RESULT 13
US-10-282-122A-55748
; Sequence 55748, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A

```

```

; SOFTWARE: Patentin version 3.1
; SEQ ID NO 67145
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-282-122A-67145

Query Match      44.0%; Score 117; DB 15; Length 389;
Best Local Similarity 55.9%; Pred. No. 0.002;
Matches 33; Conservative 7; Mismatches 15; Indels 4; Gaps 2;

QY   1 AKKYAKKEKAYAK--KAEEAAKKAEKAYKAAEAKKAEKAYKAAEAKKAAEAKKAAEAAVEA 56
     |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   189 AKRKAEEAKAEAKAKVKAEEAAKV-KAEKAAEAAEAKAEKAAEAKAEKAAEAKAEKAA 246

RESULT 9
US-10-282-122A-56483
; Sequence 56483, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 56483
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-56483

Query Match      44.0%; Score 117; DB 15; Length 421;
Best Local Similarity 54.7%; Pred. No. 0.0021;
Matches 35; Conservative 5; Mismatches 10; Indels 14; Gaps 2;

QY   7 KEKAYAKAEKAEKAEKAYK-----AAEKKAEEA-----KYKAAEAATAAKEA 52
     |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   148 KAEADAKAEAEKAAADAKKAEAAEAKAAEAKAAEAKAAEAKAAEAKAAEAKAAEAKAA 207

SOFTWARE: Patentin version 3.1
; SEQ ID NO 60543
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60543

Query Match      43.8%; Score 116.5; DB 15; Length 239;
Best Local Similarity 55.4%; Pred. No. 0.0013;
Matches 31; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

QY   1 AKKYAKKEKAYAKAEKAAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAAEAAVE 55
     |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   116 AEKAAEEKKAAAEKAEADKKQEEDAVKAAAKKEQAEEAAEBEKAADAKAAAEKAAAE 171

RESULT 11
US-10-282-122A-75047
; Sequence 75047, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl

```

```
QY 1 AKKYAKK-EKAYAKA-----EKAAKKAKAYKAAEAKKKAEAK----- 39
      ||||| ||||| :|| ||||| ||||| ||||| :|
Db 1 AKKYAKKAEKAYAKKAKAEKKAYAKKAEKAYKAAEAKKKAKAEKAYAKKAEAKKKEA 60
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 40 -----YKAEAKAAAEKAEAYEA 56
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 YKAEAKKYAKAAAEKKEYYAAAEAKKAEAAKAYKAAEAKAAAEKAEAYEA 109
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
US-09-816-989A-2
; Sequence 2, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-2

Query Match 50.6%; Score 134.5; DB 9; Length 45;
Best Local Similarity 68.4%; Pred. No. 4.5e-06;
Matches 39; Conservative 0; Mismatches 5; Indels 13; Gaps 4;

QY 1 AKKYAKKERAYAKKAEKAKAEKAAEAKKKAEAKYKAAEAKKKAEAKYKAAEAKKAEAYEA 56
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 AKKYAKK-----AKAEKA-----KKAYKAAEAKK--AAKYKAAAEKAAAEKAEAYEA 45
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-10-282-122A-68109
; Sequence 68109, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
```

```
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68109
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-282-122A-68109

Query Match 44.0%; Score 117; DB 15; Length 372;
Best Local Similarity 66.0%; Pred. No. 0.0019;
Matches 31; Conservative 2; Mismatches 12; Indels 2; Gaps 1;

QY 7 KEKAYAKKAEKAAKAEKAYKAAEAKKKAEAKYKAAEAKKKAEAKYKAAEAKKAEAA 53
      :|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :|
Db 113 REAAEAKKAEADAAKAAEAAA--KAAEAKKAAEAKKAEAKKAEAKKAAEKQQA 157
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-10-282-122A-67145
; Sequence 67145, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
```



Result No.	Query			Description	
	Score	Match	Length	ID	
1	266	100.0	56	9 US-09-816-989A-3	Sequence 3, Appli
2	245.5	92.3	77	9 US-09-816-989A-5	Sequence 5, Appli
3	233	87.6	86	9 US-09-816-989A-6	Sequence 6, Appli
4	198	74.4	66	9 US-09-816-989A-4	Sequence 4, Appli
5	180.5	67.9	109	9 US-09-816-989A-7	Sequence 7, Appli
6	134.5	50.6	45	9 US-09-816-989A-2	Sequence 2, Appli
7	117	44.0	372	15 US-10-282-122A-68109	Sequence 68109, A
8	117	44.0	389	15 US-10-282-122A-67145	Sequence 67145, A
9	117	44.0	421	15 US-10-282-122A-56483	Sequence 56483, A
10	116.5	43.8	239	15 US-10-282-122A-60543	Sequence 60543, A
11	113.5	42.7	407	15 US-10-282-122A-75047	Sequence 75047, A
12	112.5	42.3	387	15 US-10-282-122A-72645	Sequence 72645, A
13	111	41.7	428	15 US-10-282-122A-55748	Sequence 55748, A

## RESULT 2

Search completed: December 14, 2004, 06:01:12  
Job time : 62.9072 secs

---













```

XX OS Unidentified.
XX KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;
XX KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; antipruritic; dermatological;
XX KW antianaemic; immunosuppressive; demyelinating disease; osteoarthritis;
XX PN inflammatory condition; multiple sclerosis; rheumatoid arthritis;
XX PD Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;
XX KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;
XX KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;
XX KW pemphigus vulgaris; systemic lupus erythematosus.
XX OS Unidentified.
XX PN WO200018794-A1.
XX PD 06-APR-2000.
XX PF 24-SEP-1999; 99WO-US022402.
XX PR 25-SEP-1998; 98US-0101693P.
XX PA (YEDA ) YEDA RES & DEV CO LTD.
XX PA (TEVA-) TEVA PHARM USA INC.
XX PI Gad A, Lis D;
XX WPI; 2000-317499/27.
XX Copolymer 1 related polypeptides used as molecular weight markers for
XX glatiramer acetate and for treatment and prevention of immune diseases.
XX Claim 10; Page 14; 72pp; English.
XX AAY82571 to AAY82577 represent specifically claimed copolymer molecular
XX weight TV-marker polypeptides from the present invention. The present
XX invention describes polypeptides (I) for determining the molecular weight
XX of a copolymer (CP), which has an identified molecular weight and an
XX amino acid composition corresponding to the copolymer. The polypeptides
XX of the invention are used as molecular weight markers for glatiramer
XX acetate related tetrapolymers. The polypeptides may also be used for
XX treating and preventing immune diseases in a mammal. Autoimmune diseases
XX which may be treated include either cell-mediated or antibody-mediated
XX diseases. Such diseases include arthritic conditions, demyelinating
XX diseases and inflammatory conditions, e.g. multiple sclerosis, rheumatoid
XX arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune
XX oophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's
XX disease, chronic immune thrombocytopaenia purpura, colitis, contact
XX sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's
XX syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,
XX psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-
XX diseases which can be treated include host-versus-graft disease,
XX graft-versus-host disease, and delayed-type hypersensitivity. The
XX polypeptides of the invention have defined molecular weights and physical
XX properties which are analogous to glatiramer acetate molecules, which
XX makes them ideal for use as molecular weight markers
XX Sequence 109 AA;
XX
XX Query Match 67.9%; Score 180.5; DB 3; Length 109;
XX Best Local Similarity 45.9%; Pred. No. 5.8e-11;
XX Matches 50; Conservative 3; Mismatches 3; Indels 53; Gaps 3;
XX
XX QY 1 AKKYAKK-EKAYAKKA-----EKAANKKAEKAYKAAEAKKAEAK----- 39
XX Db 1 AKKYAKKAEKAYAKKAKEKAYAKKAKEKAYAKKAKEKAYAKKAKEKAYAKKAKEA 60
XX
XX QY 40 -----YKAEAAKAAKAAKEAAAYEA 56
XX Db 61 YKAEAKKYAKAAKAEKAYAAEAKKAAKAAKAYKAAEAKKAAKAAKEAAAYEA 109
XX
XX RESULT 6
XX AAY82572
XX ID AAY82572 standard; peptide; 45 AA.
XX AC AAY82572;
XX XX
XX XX 28-JUL-2000 (first entry)
XX XX Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:2.
XX DE Copolymer; molecular weight marker; TV-marker; immune disease;
XX KW glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;

```

DR WPI; 2000-317499/27.

XX XX  
XX XX  
PT PT Copolymer 1 related polypeptides used as molecular weight markers for  
PT glatiramer acetate and for treatment and prevention of immune diseases.  
XX  
PS Claim 10; Page 14; 72pp; English.

XX CC AAY82571 to AAY82577 represent specifically claimed copolymer molecular  
weight TV-marker polypeptides from the present invention. The present  
invention describes polypeptides (I) for determining the molecular weight  
of a copolymer (CP), which has an identified molecular weight and an  
amino acid composition corresponding to the copolymer. The polypeptides  
of the invention are used as molecular weight markers for glatiramer  
acetate related tetrapolymers. The polypeptides may also be used for  
treating and preventing immune diseases in a mammal. Autoimmune diseases  
which may be treated include either cell-mediated or antibody-mediated  
diseases. Such diseases include arthritic conditions, demyelinating  
arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune  
ophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's  
disease, chronic immune thrombocytopaenia purpura, colitis, contact  
sensitvity disease, diabetes mellitus, Graves disease, Guillain-Barre's  
syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,  
psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-  
mediated diseases which can be treated include host-versus-graft disease,  
graft-versus-host disease, and delayed-type hypersensitivity. The  
polypeptides of the invention have defined molecular weights and physical  
properties which are analogous to glatiramer acetate molecules, which  
makes them ideal for use as molecular weight markers

XX SQ Sequence 86 AA;

Query Match 87.6%; Score 233; DB 3; Length 86;  
Best Local Similarity 62.8%; Pred. No. 2.3e-16;  
Matches 54; Conservative 2; Mismatches 0; Indels 30; Gaps 1;

OY 1 AKKYAKKEKAYAKAEEAAKAEAYKAABAKKKAAEAK----- 39  
| | | | | | | | | | | | | | | | | | | | : |  
DB 1 AKKYAKKEKAYAKAEEAAKAEAYKAABAKKKAAEAKKKYAKAAEKKEYAAEA 60

OY 40 -----YKAEAAKAAKEAAEYA 56  
| | | | | | | | | | | | | | | | | | | | : |  
DB 61 YKAEAAKAYKAEEAAKAAEAYKAABAKKKAAEAKKKYAKAAEKKEYAAEA 86

RESULT 4  
AAY82574  
ID AAY82574 standard; peptide; 66 AA.  
AC AAY82574;  
XX  
DT 28-JUL-2000 (first entry)  
XX

Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:4.

DE DE  
XX Copolymer; molecular weight marker; TV-marker; immune disease;  
KW glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;  
KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;  
KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;  
KW antianaemic; immunosuppressive; demyelinating disease; osteoarthritis;  
KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;  
KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;  
KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;  
KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;  
KW pemphigus vulgaris; systemic lupus erythematosus.

XX Unidentified.  
OS WO200018794-A1.  
PN  
PD 06-APR-2000.  
PF 24-SEP-1999; 99WO-USO22402.

24-SEP-1999; 99WO-US022402.





GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 61.9072 Seconds  
(without alignments)  
324.499 Million cell updates/sec

Title: US-10-792-311-3  
Perfect score: 266  
Sequence: 1 AKKYAKKEKAVAKAEKAAK.....EAKYKAEAKAAKEAAVEA 56

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04: \*  
1: Geneseqp1980s: \*  
2: Geneseqp1990s: \*  
3: Geneseqp2000s: \*  
4: Geneseqp2001s: \*  
5: Geneseqp2002s: \*  
6: Geneseqp2003as: \*  
7: Geneseqp2003bs: \*  
8: Geneseqp2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	266	100.0	56	3	AAY82573	Copolymer
2	245.5	92.3	77	3	AAV82575	Copolymer
3	233	87.6	86	3	AAV82576	Copolymer
4	198	74.4	66	3	AAV82574	Copolymer
5	180.5	67.9	109	3	AAV82577	Copolymer
6	134.5	50.6	45	3	AAV82572	Copolymer
7	117	44.0	372	6	ABU40185	Protein e
8	117	44.0	389	6	ABU39221	Protein e
9	117	44.0	421	6	ABU28559	Protein e
10	116.5	43.8	239	5	ABU49123	Listeria
11	116.5	43.8	239	6	ABU32619	Protein e
12	113.5	42.7	407	6	ABU47123	Protein e
13	112.5	42.3	387	6	ABU44721	Protein e
14	111	41.7	106	2	AAR06446	Recombina
15	111	41.7	428	6	ABU27824	Protein e
16	110.5	41.5	376	6	ABU47848	Protein e
17	108.5	40.8	323	6	ABU31397	Protein e
18	108.5	40.8	469	7	ABO67048	Klebsiell
19	105.5	39.7	372	5	ABG80418	Haemophil
20	105.5	39.7	372	7	ABO23507	Haemophil
21	104.5	39.3	35	3	AAV82571	Copolymer
22	104	39.1	100	3	AAV98499	Peptide #
23	104	39.1	100	3	AAV59044	Amino aci
24	104	39.1	100	4	AAV45852	Nucleic a
25	104	39.1	100	4	AAU04289	Poly-Lys-

26	103	38.7	154	2	AAR06445	Recombina
27	103	38.7	223	2	AAV14928	Amino aci
28	103	38.7	223	6	ABP70903	Mycobacte
29	103	38.7	361	7	ADF05105	Bacterial
30	102.5	38.5	347	6	ABU38313	Protein e
31	102.5	38.5	347	6	ABJ18771	Pseudomon
32	102.5	38.5	407	7	ABO80835	Pseudomon
33	102	38.3	146	4	AAV25508	Human pro
34	102	38.3	357	6	ABM67869	Photorhab
35	101	38.0	1507	3	AAV24128	Plasmodiu
36	100	37.6	214	6	ADA33882	Acinetoba
37	100	37.6	700	7	ADC95467	E. faeciu
38	99.5	37.4	80	5	ABG71044	Tumour ne
39	98.5	37.0	336	6	ABU42038	Protein e
40	98	36.8	60	5	ABG71043	Tumour ne
41	98	36.8	427	4	AAV70868	C albican
42	98	36.8	427	8	ADP99051	Adp99051
43	97.5	36.7	205	3	AAV20575	Mycobacte
44	96.5	36.3	452	6	ABP57088	Mouse MKP
45	95	35.7	78	7	ADE10604	Structura

ALIGNMENTS

RESULT 1

AAV82573  
ID AAY82573 standard; peptide; 56 AA.

XX	AC	AAV82573;
XX	DT	28-JUL-2000 (first entry)
XX	DE	Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:3.
XX	XX	Copolymer; molecular weight marker; TV-marker; immune disease; glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective; osteopathic; immunosuppressive; antithyroid; antiinflammatory; antidiabetic; thyronimetic; haemostatic; antipsoriatic; dermatological; antianemic; immunosuppressive; demyelinating disease; osteoarthritis; inflammatory condition; multiple sclerosis; rheumatoid arthritis; Crohn's disease; chronic immune thrombocytopaenia purpura; colitis; diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis; Hashimoto's disease; idiopathic myxoedema; myasthenia gravis; pemphigus vulgaris; systemic lupus erythematosus.
XX	XX	Unidentified.
OS	XX	WO200018794-A1.
PN	XX	06-APR-2000.
PD	XX	24-SEP-1999; 99WO-US022402.
PF	XX	25-SEP-1999; 98US-0101693P.
PR	XX	(YEDA ) YEDA RES & DEV CO LTD.
PA	XX	(TEVA-) TEVA PHARM USA INC.
PI	XX	Gad A, Lis D;
PI	XX	WPI; 2000-317499/27.
DR	XX	Copolymer 1 related polypeptides used as molecular weight markers for glatiramer acetate and for treatment and prevention of immune diseases.
PT	XX	Claim 10; Page 14; 72pp; English.
PT	XX	AAV82571 to AAV82577 represent specifically claimed copolymer molecular weight TV-marker polypeptides from the present invention. The present invention describes polypeptides (I) for determining the molecular weight of a copolymer (CP), which has an identified molecular weight and an amino acid composition corresponding to the copolymer. The polypeptides



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 72.962 Seconds  
(without alignments)  
324.499 Million cell updates/sec

Title: US-10-792-311-4  
Perfect score: 313  
Sequence: 1 AKYAKKEKAYAKKAEAK.....EAKYKAEAKAAKAEAYEA 66

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	313	100.0	66	3 AAY82574	Aay82574 Copolymer
2	282.5	90.3	77	3 AAY82575	Aay82575 Copolymer
3	268	85.6	86	3 AAY82576	Aay82576 Copolymer
4	228.5	73.0	109	3 AAY82577	Aay82577 Copolymer
5	198	63.3	56	3 AAY82578	Aay82578 Copolymer
6	131	41.9	421	6 ABU28559	Abu28559 Protein e
7	130	41.5	323	6 ABU31397	Abu31397 Protein e
8	130	41.5	428	7 ABU27824	Abu27824 Protein e
9	130	41.5	469	7 ABU67048	Abu67048 Klebsiell
10	129.5	41.4	154	2 AAR06445	Aar06445 Recombina
11	128	40.9	100	3 AAY98499	Aay98499 Peptide #
12	128	40.9	100	3 AAY59044	Aay59044 Amino aci
13	128	40.9	100	4 AAB45852	Aab45852 Nucleic a
14	128	40.9	100	4 AAU04289	Aau04289 Poly-Lys-
15	126.5	40.4	389	6 ABU39221	Abu39221 Protein e
16	124.5	39.8	407	6 ABU47123	Abu47123 Protein e
17	122.5	39.1	372	6 ABU40185	Abu40185 Protein e
18	122.5	39.1	376	6 ABU47848	Abu47848 Protein e
19	122	39.0	106	2 AAR06446	Aar06446 Recombina
20	121	38.7	387	6 ABU44721	Abu44721 Protein e
21	120.5	38.5	45	3 AAY82572	Aay82572 Copolymer
22	119.5	38.2	347	6 ABU38313	Abu38313 Protein e
23	119.5	38.2	347	6 ABU18771	Abj18771 Pseudomon
24	119.5	38.2	407	7 AB080835	Ab080835 Pseudomon
25	117.5	37.5	372	5 ABG80418	Abg80418 Haemophil

26	117.5	37.5	372	7 ABO23507	Ab023507 Haemophil
27	117	37.4	357	6 ABM67869	Abm67869 Photorhab
28	117	37.4	388	6 ABU50266	Abu50266 Protein e
29	115	36.7	336	6 ABU42038	Abu42038 Protein e
30	114.5	36.6	361	7 ADF05105	Adf05105 Bacterial
31	114	36.4	452	6 ABP57088	Abp57088 Mouse MKP
32	113.5	36.3	214	6 ADA33882	Ada33882 Acinetoba
33	111.5	35.6	239	5 ABB49123	Abb49123 Listeria
34	111.5	35.6	239	6 ABU32619	Abu32619 Protein e
35	110.5	35.3	205	3 AAB20575	Aab20575 Mycobacte
36	109.5	35.0	80	5 ABG71044	Abg71044 Tumour ne
37	108.5	34.7	198	4 AAB95499	Ab095499 Human pro
38	108.5	34.7	279	4 AAU03592	Aau03592 Human DNA
39	108.5	34.7	467	4 AAB94309	Abp94309 Human pol
40	108.5	34.7	467	5 ABP69558	Abp69558 Human pol
41	106	33.9	214	2 AAY34055	Aay34055 M. tuberc
42	106	33.9	214	2 AAY57353	Aay57353 M. tuberc
43	106	33.9	214	6 ABU34623	Abu34623 Protein e
44	106	33.9	214	6 ABU36893	Abu36893 Protein e
45	105.5	33.7	35	3 AAY82571	Aay82571 Copolymer

ALIGNMENTS

RESULT 1  
AAY82574  
ID AAY82574 standard; peptide; 66 AA.

XX AAY82574;

XX 28-JUL-2000 (first entry)

XX Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:4.

Copolymer; molecular weight marker; TV-marker; immune disease;  
glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;  
osteopathic; immunosuppressive; antithyroid; antiinflammatory;  
antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;  
antianemic; immunosuppressive; demyelinating disease; osteoarthritis;  
inflammatory condition; multiple sclerosis; rheumatoid arthritis;  
Crohn's disease; chronic immune thrombocytopenia purpura; colitis;  
diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;  
Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;  
pemphigus vulgaris; systemic lupus erythematosus.

XX Unidentified.

XX WO200018794-A1.

XX 06-APR-2000.

XX 24-SEP-1999; 99WO-US022402.

XX 25-SEP-1998; 98US-0101693P.

XX (YEDA ) YEDA RES & DEV CO LTD.  
(TEVA-) TEVA PHARM USA INC.

XX Gad A, Lis D;

XX WPI; 2000-317499/27.

Copolymer 1 related polypeptides used as molecular weight markers for  
glatiramer acetate and for treatment and prevention of immune diseases.

Claim 10; Page 14; 72pp; English.

AAY82571 to AAY82577 represent specifically claimed copolymer molecular  
weight TV-marker polypeptides from the present invention. The present  
invention describes polypeptides (I) for determining the molecular weight  
of a copolymer (CP), which has an identified molecular weight and an  
amino acid composition corresponding to the copolymer. The polypeptides



[illegible]











```
Query Match      40.9%; Score 128; DB 3; Length 100;
Best Local Similarity 62.7%; Pred. No. 2.1e-05;
Matches 42; Conservative 5; Mismatches 16; Indels 4; Gaps 4;

QY 1 AKKYAK-KEKAYAK-KAEAKAAKAKAEAKKYAKA-AKAEKKEVAAAEAKYKAEA-AK 56
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 AKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAK 61

QY 57 AAKEAA 63
   ||| |||
Db 62 AKAKAKA 68

RESULT 13
AAB45852
ID AAB45852 standard; protein; 100 AA.
XX
AC AAB45852;
XX
DT 21-MAR-2001 (first entry)
XX
DE Nucleic acid transporter system peptide ligand SEQ ID NO 64.
XX
KW Nucleic acid delivery; nucleic acid transporter system; hormone; enzyme;
KW growth factor; clotting factor; apolipoprotein; receptor; drug; oncogene;
KW tumor antigen; tumor suppressor; viral antigen; parasitic antigen;
KW bacterial antigen.
XX
OS Unidentified.
XX
PN US6150168-A.
XX
PD 21-NOV-2000.
XX
PF 05-JUN-1995; 95US-00460971.
XX
PR 20-MAR-1992; 92US-00855389.
PR 19-MAR-1993; 93MO-US002725.
PR 14-DEC-1993; 93US-00167641.
XX
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI Gottchalk S, Sparrow J, Cristiano RJ, Smith LC, Woo SLC;
XX WPI; 2001-049093/06.
XX
PT Nucleic acid transporter system for delivering nucleic acid into a cell,
PT useful for delivering proteins and polypeptides to cells, including
PT growth factors, enzymes, hormones, and tumor suppressors.
XX
PS Disclosure; Col 125-126; 105pp; English.
XX
CC This invention describes a novel system (I) for delivering a nucleic acid
CC to a cell, comprising a binding complex comprising a ligand binding
CC molecule noncovalently bound to a nucleic acid and covalently linked to a
CC surface ligand, and a second binding complex comprising a second binding
CC molecule noncovalently bound to a nucleic acid and covalently linked to a
CC nuclear ligand. The complexes are simultaneously bound to the nucleic
CC acid. The nucleic acid transporter system can also be used in a method
CC for the in vivo targeting of the insertion of DNA into a cell. It can
CC also be used in processes for producing transformed cell lines. The
CC system can be used to deliver a variety of proteins and polypeptides,
CC such as hormones, growth factors, enzymes, clotting factors,
CC apolipoproteins, receptors, drugs, oncogenes, tumor antigens, tumor
CC suppressors, viral antigens, parasitic antigens, and bacterial antigens.
CC The transporter system uses lysis agents to overcome the problems of
CC endosomal/lysosomal degradation seen with prior art systems
XX
SQ Sequence 100 AA;

Query Match      40.9%; Score 128; DB 4; Length 100;
Best Local Similarity 62.7%; Pred. No. 2.1e-05;
Matches 42; Conservative 5; Mismatches 16; Indels 4; Gaps 4;

QY 1 AKKYAK-KEKAYAK-KAEAKAAKAKAEAKKYAKA-AKAEKKEVAAAEAKYKAEA-AK 56
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 AKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAK 61

QY 57 AAKEAA 63
   ||| |||
Db 62 AKAKAKA 68

RESULT 14
AAU04289
ID AAU04289 standard; peptide; 100 AA.
XX
AC AAU04289;
XX
DT 23-OCT-2001 (first entry)
XX
DE Poly-Lys-Ala used in nucleic acid transporter system.
XX
KW Nucleic acid transport; cytolysis; ligand; lysis agent; spacer molecule;
KW gene therapy; hepatocyte; muscle; bone forming cell.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 3..100
FT /note="Lys-Ala in positions 3-100 may be present or
FT absent"
XX
PN US6177554-B1.
XX
PD 23-JAN-2001.
XX
PF 05-JUN-1995; 95US-00462040.
XX
PR 20-MAR-1992; 92US-00855389.
PR 19-MAR-1993; 93MO-US002725.
PR 14-DEC-1993; 93US-00167641.
XX
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI Woo SLC, Smith LC, Cristiano RJ, Gottchalk S, Sparrow J;
XX WPI; 2001-365933/38.
XX
PT Nucleic acid transport system, useful for creating transgenic animals for
PT assessing human disease such as cancer in an animal model.
XX
PS Disclosure; Col 131; 111pp; English.
XX
CC The sequence represents poly-Lys-Ala, used to bind nucleic acid in a
CC nucleic acid transporter system. The nucleic acid transporter system uses
CC nucleic acid binding complexes containing surface ligands which are
CC capable of binding to a cell surface receptor and entering the cell
CC through cytolysis. The compounds of the invention are either ligands,
CC binding molecules (surface ligands), lysis agents, spacer molecules or
CC their intermediates. The ligands, binding molecules, lysis agents and
CC spacer molecules are used in nucleic acid transporter systems to deliver
CC nucleic acid into specific cells e.g. in gene therapy to deliver nucleic
CC acid into hepatocytes, muscle cells or bone forming cells
XX
SQ Sequence 100 AA;

Query Match      40.9%; Score 128; DB 4; Length 100;
Best Local Similarity 62.7%; Pred. No. 2.1e-05;
Matches 42; Conservative 5; Mismatches 16; Indels 4; Gaps 4;

QY 1 AKKYAK-KEKAYAK-KAEAKAAKAKAEAKKYAKA-AKAEKKEVAAAEAKYKAEA-AK 56
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 AKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAK 61
```



QY 57 AAKEAA 63  
Db 62 AKAKAKA 68

RESULT 15

ABU39221  
ID ABU39221 standard; protein; 389 AA.

XX AC ABU39221;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #24748.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Pasteurella multocida.

XX PN WO20027183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR N-PSDB; ACA43091.

XX PT New antisense nucleic acids, useful for identifying proteins or screening  
XX PT for homologous nucleic acids required for cellular proliferation to  
XX PT isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 67145; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of  
XX CC the 6213 antisense sequences given in the specification where expression  
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid  
XX CC encoding a polypeptide whose expression is inhibited by the antisense  
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX CC polypeptide or its fragment whose expression is inhibited by the  
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding  
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX CC proliferation or the activity of a gene in an operon required for  
XX CC proliferation; (7) identifying a compound that influences the activity of  
XX CC the gene product or that has an activity against a biological pathway  
XX CC required for proliferation, or that inhibits cellular proliferation; (8)  
XX CC identifying a gene required for cellular proliferation or the biological  
XX CC pathway in which a proliferation-required gene or its gene product lies  
XX CC or a gene on which the test compound that inhibits proliferation of an  
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX CC compound's activity; (11) a culture comprising strains in which the gene  
XX CC product is overexpressed or underexpressed; (12) determining the extent  
XX CC to which each of the strains is present in a culture or collection of  
XX CC strains; or (13) identifying the target of a compound that inhibits the  
XX CC proliferation of an organism. The antisense nucleic acids are useful for  
XX CC identifying proteins or screening for homologous nucleic acids required  
XX CC for cellular proliferation to isolate candidate molecules for rational  
XX CC drug discovery programs, or for screening homologous nucleic acids  
XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
XX CC *K. pneumoniae* or *F. aeruginosa*. The present sequence is encoded by one of  
XX CC the target prokaryotic essential genes. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 389 AA;

Query Match 40.4%; Score 126.5; DB 6; Length 389;  
Best Local Similarity 52.0%; Pred. No. 0.00012;  
Matches 39; Conservative 9; Mismatches 12; Indels 15; Gaps 3;  
QY 1 AKKYAKKEKAVAKAK-----KAEAKA---AKKAKAEAKKYAKAKAEKKEKAYAAAEAKYK 51  
Db 189 AKRKAEEKAKAEAEAKAKVEKAKAEAEAKVKAEAKAEAEAKAKAEKAK-----AEAKAK 242  
QY 52 AFAAKAAAKEAAYEA 66  
Db 243 AEKAKADAEAAQORKA 257

Search completed: December 14, 2004, 06:01:13  
Job time : 73.962 secs

***This Page Blank (uspto)***

Qy 61 EAAVEA 66  
| | | | |







RESULT 11

US-10-282-122A-68109

Sequence 68109, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

RESULT 12  
US-10-282-122A-75772  
; Sequence 75772, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.

```

1  APPLICATION: AU, H.
2  TITLE INVENTION: Identification of Essential Genes in Microorganisms
3  FILE REFERENCE: ELITRA.034A
4  CURRENT APPLICATION NUMBER: US/10/282,122A
5  CURRENT FILING DATE: 2003-02-20
6  PRIOR APPLICATION NUMBER: 60/191,078
7  PRIOR FILING DATE: 2000-03-21
8  PRIOR APPLICATION NUMBER: 60/206,848
9  PRIOR FILING DATE: 2000-05-23
10 PRIOR APPLICATION NUMBER: 60/207,727
11 PRIOR FILING DATE: 2000-05-26
12 PRIOR APPLICATION NUMBER: 60/230,335
13 PRIOR FILING DATE: 2000-09-06
14 PRIOR APPLICATION NUMBER: 60/230,347
15 PRIOR FILING DATE: 2000-09-09
16 PRIOR APPLICATION NUMBER: 60/242,578
17 PRIOR FILING DATE: 2000-10-23
18 PRIOR APPLICATION NUMBER: 60/253,625
19 PRIOR FILING DATE: 2000-11-27
20 PRIOR APPLICATION NUMBER: 60/257,931
21 PRIOR FILING DATE: 2000-12-22
22 PRIOR APPLICATION NUMBER: 60/267,636
23 PRIOR FILING DATE: 2001-02-09
24 PRIOR APPLICATION NUMBER: 60/269,308
25 PRIOR FILING DATE: 2001-02-16
26 Remaining Prior Application data removed - See File Wrapper or PALM.
27 NUMBER OF SEQ ID NOS: 78614
28 SOFTWARE: PatentIn version 3.1
29 SEQ ID NO 75772
30 LENGTH: 376
31 TYPE: PRT

```







**This Page Blank (uspto)**









```

; FILE REFERENCE: 60807-A
; CURRENT APPLICATION NUMBER: US/09/405,743A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
US-09-405-743A-2

Query Match      38.5%; Score 120.5; DB 4; Length 45;
Best Local Similarity 71.1%; Pred. No. 2.3e-05;
Matches 33; Conservative 0; Mismatches 6; Indels 7; Gaps 2;

QY      29 AKKYAKAAKAE--KKEYAARAEK-----YKAEAKAKAAAEAAVEA 66
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      1 AKKYAKKAERAKKAYKAAAEKKAAYEKAAAEKKAARAEAAVEA 45

RESULT 12
US-09-252-991A-29581
; Sequence 29581, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29581
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29581

Query Match      38.2%; Score 119.5; DB 4; Length 407;
Best Local Similarity 51.5%; Pred. No. 0.00027;
Matches 34; Conservative 7; Mismatches 22; Indels 3; Gaps 1;

QY      1 AKKYAK---KEKAYAKAKAEAKAAKAAKAAEKYAKAAKAAEKYAAAEEAKYKAAEAA 57
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      215 AKKRADEAKKAAEDAKKAAEDAKKAAAEAKKAAAEAAVEAAKKKAAEAA 274

QY      58 AAKEAA 63
      ||::||
DB      275 AAKAA 280

RESULT 13
US-08-993-008A-6
; Sequence 6, Application US/08993008A
; Patent No. 6153596
; GENERAL INFORMATION:
; APPLICANT: Liotta, Dennis C.
; APPLICANT: Petros, John A.
; APPLICANT: Wey, Shlow-Jyi
; APPLICANT: Karr, Joan F.
; APPLICANT: Pohl, Jan
; TITLE OF INVENTION: Polycationic Oligomers
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan
; STREET: 5370 Manhattan Circle, Suite 201

```

Db 233 KAAQQ 238

RESULT 15

US-09-328-352-5169

; Sequence 5169, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 5169

; LENGTH: 214

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-5169

Query Match 36.3%; Score 113.5; DB 4; Length 214;

Best Local Similarity 51.4%; Pred. No. 0.00052;

Matches 37; Conservative 9; Mismatches 17; Indels 9; Gaps 4;

QY 1 AKKYAKKEKAYAKAKAEAKKAKAEAKKYAKAAK--AEKKEYA-----AAEAKYKAE 53

Db 55 AKRQAEADKKAAEAKR-QAEADKKA-AEAKRQAEADKKAAEAKRQAEADKKAAEAKRKA 112

QY 54 AAQAAAEAAAYE 65

Db 113 AEKKAEEAKARE 124

Search completed: December 14, 2004, 05:50:17

Job time : 18.6618 secs





```
membrane spanning protein Tola [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F90725
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.-G.
gaawara, N.; Yaenaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90725
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-394 <HAY>
A;Cross-references: UNIPROT:Q8X965; GB:BA000007; PIDN:BAB34197.1; PID:g13360233; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 050952
C;Genetics:
A;Gene: EC80774

Query Match 40.4%; Score 148; DB 2; Length 394;
Best Local Similarity 55.1%; Pred. No. 0.00033;
Matches 49; Conservative 7; Mismatches 21; Indels 12; Gaps 4;

QY 1 AKKYAKKEKAYA-----KKAEEAKKAEAKYKAAEAK-KKAKAEAKKYAKAA-----KA 49
DB 138 AAKAAADAKAKAADDKAAEEAKKAAADAKKKAEEAKKAAAEAKKKAEEAKKKAEEAKKKA 197

QY 50 EKKEYAAAEAKYKAAKAAA-KEAAVEA 77
DB 198 EAAEAAAAEAKKAAAEAKKAAAEAKKAAAEKA 226

RESULT 3
G85576
membrane spanning protein Tola [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: G85576
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: G85576
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-394 <STO>
A;Cross-references: UNIPROT:Q8X965; GB:AE005174; NID:g12513672; PIDN:AAG55075.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: tola

Query Match 40.4%; Score 148; DB 2; Length 394;
Best Local Similarity 55.1%; Pred. No. 0.00033;
Matches 49; Conservative 7; Mismatches 21; Indels 12; Gaps 4;

QY 1 AKKYAKKEKAYA-----KKAEEAKKAEAKYKAAEAK-KKAKAEAKKYAKAA-----KA 49
DB 138 AAKAAADAKAKAADDKAAEEAKKAAADAKKKAEEAKKAAAEAKKKAEEAKKKAEEAKKKA 197

QY 50 EKKEYAAAEAKYKAAKAAA-KEAAVEA 77
DB 198 EAAEAAAAEAKKAAAEAKKAAAEAKKAAAEKA 226

RESULT 4
AG0592
tola protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain Ctr18
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AG0592
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
```

```
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Ouail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AG0592
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-376 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05209.1; PID:g16501979; GSPDB:GN00176
C;Genetics:
A;Gene: STY0793

Query Match 39.9%; Score 146; DB 2; Length 376;
Best Local Similarity 54.4%; Pred. No. 0.00044;
Matches 49; Conservative 7; Mismatches 14; Indels 20; Gaps 5;

QY 1 AKKYAKKEKAYA-KAAKAAKAAKAYK-AAEAKKAKAKAEAKKYAKAAKAEKK-----EY 54
DB 158 AKKKAEEA--AKAAADAKKKAEEA-KAAAEAKKAAEAAK--AAADAKKKADAEEAAK 213

QY 55 AAEAKYKAEAA-----KAAAEAA 74
DB 214 AAEEAKKKADAAAKAAADAKKKAEEAA 243

RESULT 5
E83525
TolA protein PA0971 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: E83525
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Brj
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83525
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-347 <STO>
A;Cross-references: UNIPROT:P50600; GB:AE004530; GB:AE004091; NID:g9946865; PIDN:AAG0436
A;Experimental source: strain PA01
C;Genetics:
A;Gene: tola; PA0971

Query Match 38.7%; Score 141.5; DB 2; Length 347;
Best Local Similarity 44.1%; Pred. No. 0.00089;
Matches 45; Conservative 9; Mismatches 19; Indels 29; Gaps 3;

QY 1 AKKYAKKEKAYA-KAAKAAKAEAKA-----YKAAE-A 32
DB 120 ARKAAEQAAEAKKAAEAKKAAEAKAAEQKQADIAKKAEDAEAKKAAEDAKKKAEDA 179

QY 33 KKKAKAEAKYKAAKAEKKEKAAEAKYKAAEAKKAAEAA 74
DB 180 KKKAAEEAKKA-AAEAKKKAEEAKKAAEAKKAAEAA 220

RESULT 6
G64064
outer membrane integrity protein tola - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C;Accession: G64064; JC5212
R;Gleitschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
```

A:Accession: G64064  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-372 <TRIG>  
A:Cross-references: UNIPROT:P44678; GB:U32722; GB:L42023; NID:g1573348; PIDN:AAC22041.1;  
R:Sen, K.; Sikkema, D.J.; Murphy, T.F.  
Gene 178, 75-81, 1996  
A:Title: Isolation and characterization of the Haemophilus influenzae tolQ, tolR, tolA  
A:Accession number: JCS212; MUID:97080550; PMID:8921895  
A:Accession: JCS212  
A:Molecule type: DNA  
A:Status: preliminary  
A:Residues: 'V', 2-47, 'A', 49-141, 'R', 143-164, 'P', 166-189, 'R', 191-202, 'A', 204-226, 'A', 228-  
A:Cross-references: GB:U32470  
A:Experimental source: strain 1479  
A:Note: the authors translated the codon CGT for residue 190 as Ala  
C:Genetics:  
A:Gene: tolA  
A:Start codon: GTG

Query Match 38.5%; Score 141; DB 2; Length 372;  
Best Local Similarity 54.7%; Pred. No. 0.001;  
Matches 47; Conservative 7; Mismatches 20; Indels 12; Gaps 4;

QY 1 AKYA-----KKEKAYAKAEAAK-----AEAKYKAEAKKAKAEAKKYAKAKAEK 51  
DB 158 AKRLAAAKQAEFEAEAKAKAEIAAQKQAEAKAKLEAEAKAKAVAEAK--AK-AEAEA 214  
QY 52 KEYAAAEAKYKAEAKAAKAAKEAYEA 77  
DB 215 KAKAAAEAKAKADAEEAKATEAKRKA 240

RESULT 7  
T09127  
probable erythrocyte-binding protein MAEBL - Plasmodium yoelii  
C:Species: Plasmodium yoelii  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C:Accession: T09127  
R:Kapke, S.H.I.; Noe, A.R.; Fraser, T.S.; Blair, P.L.; Adams, J.H.  
Proc. Natl. Acad. Sci. U.S.A. 95, 1230-1235, 1998  
A:Title: A family of chimeric erythrocyte binding proteins of malaria parasites.  
A:Reference number: Z16577; MUID:98115903; PMID:9448314  
A:Accession: T09127  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1701 <KAP>  
A:Cross-references: UNIPROT:O61164; EMBL:AF031886; NID:g2947227; PID:g2947228  
A:Experimental source: subspecies yoelii; strain fM  
C:Genetics:  
A:Gene: maeb1  
A:Introns: 62/1; 1648/1; 1674/2; 1697/1  
C:Keywords: alternative splicing; cell binding; erythrocyte invasion

Query Match 37.7%; Score 138; DB 2; Length 1701;  
Best Local Similarity 54.4%; Pred. No. 0.0053;  
Matches 43; Conservative 8; Mismatches 16; Indels 12; Gaps 4;

QY 5 AKKAYAKAEKAAKAEKAAKAEAKKAKA-----EAKYAKAA-KAEKKEYA--- 55  
DB 1224 AKKAEAAKAEERKKAEKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 1282

QY 56 --AAEAKYKAEAAKAAKE 72  
DB 1283 KKAEEKKKAAEAAKAAEE 1301

RESULT 8  
A25550  
histone H1 - sea urchin (Lytechinus pictus)  
C:Species: Lytechinus pictus (painted urchin)  
C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 09-Jul-2004  
C:Accession: A25550

R:Knowles, J.A.; Childs, G.J.  
Nucleic Acids Res. 14, 8121-8133, 1986  
A:Title: Comparison of the late H1 histone genes of the sea urchins Lytechinus pictus and  
A:Reference number: A25550; MUID:87040778; PMID:3022245  
A:Accession: A25550  
A:Molecule type: DNA  
A:Residues: 1-210 <KNO>  
A:Cross-references: UNIPROT:P06144; GB:X04488; NID:g9616; PIDN:CAA28177.1; PID:g9617  
C:Superfamily: histone H1  
C:Keywords: chromosomal protein; DNA binding; nucleosome; nucleus

Query Match 37.2%; Score 136; DB 2; Length 210;  
Best Local Similarity 53.3%; Pred. No. 0.0015;  
Matches 40; Conservative 7; Mismatches 26; Indels 2; Gaps 2;

QY 2 KKYAKKERAYAKKAEKAAKAEAKYKAAEAK-KKAKAEAKKYA-KAAKAEKKEYYAAAEA 59  
DB 96 KTEAQKARAAAKKAKLAAKKKEQKKAATKARKEKLAAKKAAKAAKVKKPAKAKK 155  
QY 60 KYKAEAAKAAAKEAA 74  
DB 156 PAKKAARKPAAKKAA 170

RESULT 9  
E60110  
repetitive protein antigen 27 - Trypanosoma cruzi (fragments)  
C:Species: Trypanosoma cruzi  
C:Date: 10-Nov-1992 #sequence\_revision 10-Nov-1992 #text\_change 09-Jul-2004  
C:Accession: E60110  
R:Hofst, D.F.; Kim, K.S.; Otsu, K.; Moser, D.R.; Yost, W.J.; Blumin, J.H.; Doneleason, J.E.;  
Infect. Immun. 57, 1959-1967, 1989  
A:Title: Trypanosoma cruzi expresses diverse repetitive protein antigens.  
A:Reference number: A60110; MUID:89277508; PMID:2659529  
A:Accession: E60110  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-219 <HOF>  
A:Cross-references: UNIPROT:Q26947  
C:Comment: This protein contains a series of tandem repeats, each fourteen residues in le  
C:Superfamily: neurofilament triplet H protein  
C:Keywords: tandem repeat

Query Match 36.7%; Score 134.5; DB 2; Length 219;  
Best Local Similarity 50.7%; Pred. No. 0.002;  
Matches 38; Conservative 8; Mismatches 28; Indels 1; Gaps 1;

QY 1 AKKYAKKERAYAKKAEKAAKAEAKYKAAEAKKAEAKKAEKAYAKAK-AEKKEYYAAAEA 59  
DB 70 ATKVAGDEKQAAEATKVAEAEKQKAREATKVAEAEKQKAAEATKVAEAEKQKAAEA 129  
QY 60 KYKAEAAKAAAKEAA 74  
DB 130 TKVAEAGKQKAAEA 144

RESULT 10  
S59589  
histone H1 - Chlamydomonas reinhardtii  
C:Species: Chlamydomonas reinhardtii  
C:Date: 10-Apr-1996 #sequence\_revision 19-Apr-1996 #text\_change 09-Jul-2004  
C:Accession: S59589; S62122  
R:Fabry, S.; Mueller, K.; Lindauer, A.; Park, P.B.; Corneliuss, T.; Schmitt, R.  
Curr. Genet. 28, 333-345, 1995  
A:Title: The organization structure and regulatory elements of Chlamydomonas histone gene  
A:Reference number: S59581; MUID:96120862; PMID:8590479  
A:Accession: S59589  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-231 <FAB>  
A:Cross-references: UNIPROT:Q39576; EMBL:U16726  
A:Note: the authors did not translate the codon for residue 1  
R:Fabry, S.; Ford, C.; Lindauer, A.; Mueller, K.; Corneliuss, T.; Schmitt, R.



Best Local Similarity 51.7%; Pred. No. 0.018;  
Matches 45; Conservative 3; Mismatches 23; Indels 16; Gaps 5;  
QY 2 KKYAKKEK--AYAKAEAKK-----AEAKYAKAEAKKAKAEAKKYAKAAK-----AE 50  
DB 511 KKLAKKETAEKKCKEAKKCKEAAEKCKEAKKCKEAAEKCKEAKKCKEAKKCKEAAE 570  
QY 51 KK--EYAAAEAKYAEAAK---AAAKE 72  
DB 571 KKKCKEAKKERKEAAEKCKEAAAKE 597

RESULT 15  
G70673  
probable hupB - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: G70673  
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A: Reference number: A70500; MUID: 98295987; PMID: 9634230  
A: Accession: G70673  
A: Status: preliminary; nucleic acid sequence not shown; translation not shown  
A: Molecule type: DNA  
A: Residues: 1-214 <COL>  
A: Cross-references: UNIPROT: P95109; GB: 283018; GB: ALJ23456; NID: G3261671; PIDN: CAB05427.  
A: Experimental source: strain H37RV  
C: Genetics:  
A: Gene: hupB  
C: Superfamily: histone H1

Query Match 35.4%; Score 129.5; DB 2; Length 214;  
Best Local Similarity 45.7%; Pred. No. 0.0046;  
Matches 43; Conservative 4; Mismatches 28; Indels 19; Gaps 3;  
QY 1 AKKYAKKEKAYAKAEAKKAEAK-----AYKAAEAKKAKAEAKKAYAKAEAKKEY 54  
DB 111 AKKVAK--KAPAKKATKAAKKAATKAPARKAATKAPAKKAATKAPAKKAATKAPAKKAVATKSPAKKV 168  
QY 55 AAA-----EAKYKAEAAKAAKAEAYEA 77  
DB 169 TKAVKKTAVKASVRKKAATKAPAKKAAAKRPATKA 202

Search completed: December 14, 2004, 05:52:14  
Job time : 18.7068 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 86.4219 Seconds  
(without alignments)  
512.646 Million cell updates/sec

Title: US-10-792-311-5  
Perfect score: 366  
Sequence: 1 AKKYAKKEKAYAKAEKAAK.....EAKYKAEAAKAAKEAYEA 77

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 02:.\*  
1: uniprot\_sprot:.\*  
2: uniprot\_trembl:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	152.5	41.7	413	2	Q7C2Q4	Q7C2q4 shigella fl
2	152.5	41.7	413	2	Q83SA1	Q83sa1 shigella fl
3	151	41.3	379	2	Q7WFN5	Q7wfn5 bordetella
4	150.5	41.1	421	1	TOLA_ECOLI	P19934 escherichia
5	150	41.0	373	2	Q7W477	Q7w477 bordetella
6	148	40.4	394	2	Q7AGI8	Q7agi8 escherichia
7	148	40.4	394	2	Q8X965	Q8x965 escherichia
8	148	40.4	407	2	Q8ZQT6	Q8zqt6 salmonella
9	147.5	40.3	421	2	Q8FJT1	Q8fjt1 escherichia
10	146	39.9	376	2	Q8Z8C1	Q8z8c1 salmonella
11	145.5	39.8	395	2	Q937K4	Q937k4 erwinia chr
12	143	39.1	1701	2	Q7RC08	Q7rc08 plasmodium
13	142.5	38.9	372	2	Q9WWX1	Q9wwx1 pseudomonas
14	142.5	38.9	372	2	Q88NI6	Q88ni6 pseudomonas
15	142	38.8	395	2	Q6D7F3	Q6d7f3 erwinia car
16	141.5	38.7	347	1	TOLA_PSEAE	P50600 pseudomonas
17	141	38.5	372	1	TOLA_HAEIN	P44678 haemophilus
18	140	38.3	389	2	Q9CM70	Q9cm70 pasteurella
19	139.5	38.1	398	2	Q6MTU4	Q6mtu4 bdellovibri
20	139.5	38.1	398	2	CAE80819	Caes80819 bdellovib
21	138.5	37.8	356	2	Q7NE77	Q7net7 photorhabdu
22	138	37.7	232	2	Q39576	Q39576 chlamydomon
23	138	37.7	1652	2	Q7KEZ1	Q7kpez1 plasmodium
24	138	37.7	1652	2	AAQ73456	Aaq73456 plasmodium
25	138	37.7	1680	2	Q7KEPY9	Q7kpy9 plasmodium
26	138	37.7	1680	2	AAQ73455	Aaq73455 plasmodium
27	138	37.7	1680	2	AAQ73457	Aaq73457 plasmodium
28	138	37.7	1701	2	O61164	O61164 plasmodium
29	136	37.2	210	1	HL_LYTFI	P06144 lytechnus
30	136	37.2	713	2	Q6CDX0	Q6cdx0 yarrowia li
31	135.5	37.0	1341	2	Q98KG7	Q98kg7 rhizobium l

32	134	36.6	177	2	Q6SG84	Q6sg84 uncultured
33	134	36.6	177	2	AAR37978	Aar37978 unculture
34	134	36.6	1128	2	Q2E947	Q2e947 trypanosoma
35	132.5	36.2	401	2	Q74W64	Q74w64 yersinia pe
36	132.5	36.2	401	2	AAS61283	Aas61283 yersinia
37	131.5	35.9	105	2	Q6N503	Q6n503 rhodopseudo
38	131.5	35.9	105	2	CAE28621	Caes28621 rhodopseu
39	131.5	35.9	356	2	Q87Y39	Q87y39 pseudomonas
40	131	35.8	436	2	Q6F986	Q6f986 acinetobact
41	131	35.8	441	2	Q6N8X8	Q6n8x8 rhodopseudo
42	131	35.8	441	2	CAE27215	Caes27215 rhodopseu
43	130.5	35.7	388	2	Q8ZGZ2	Q8zgz2 yersinia pe
44	130.5	35.7	393	2	Q8CZ28	Q8cz28 yersinia pe
45	130	35.5	1391	1	MST2_DROHY	Q08696 drosophila

ALIGNMENTS

RESULT 1						
Q7C2Q4 PRELIMINARY; PRT; 413 AA.						
AC	Q7C2Q4;					
DT	05-JUL-2004 (TReMBLrel. 27, Created)					
DT	05-JUL-2004 (TReMBLrel. 27, Last sequence update)					
DT	05-JUL-2004 (TReMBLrel. 27, Last annotation update)					
DE	Membrane spanning protein.					
GN	Name=tola; OrderedLocusNames=S0571;					
OS	Shigella flexneri.					
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;					
OC	Enterobacteriaceae; Shigella.					
OX	NCBI_TaxID=623;					
RP	[1]					
RN	SEQUENCE FROM N.A.					
RC	STRAIN=2457T;					
RX	MEDLINE=22590274; PubMed=12704152;					
RA	Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,					
RA	Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,					
RA	Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,					
RA	Schwartz D.C., Blattner F.R.;					
RT	"Complete genome sequence and comparative genomics of Shigella					
RT	flexneri serotype 2a strain 2457T.";					
RL	Infect. Immun. 71:2775-2786(2003).					
DR	EMBL; AE016979; AAP16075.1; -.					
DR	InterPro; IPR010528; TOLA.					
DR	Pfam; PF06519; TOLA; 1.					
SQ	SEQUENCE 413 AA; 42355 MW; 93EI0P2C5DE60DE8 CRC64;					
Query Match 41.7%; Score 152.5; DB 2; Length 413;						
Best Local Similarity 52.4%; Pred. No. 0.00077;						
Matches 44; Conservative 11; Mismatches 22; Indels 7; Gaps 2;						
QY	1	AKKYAKKEKAYAKKAEKAAKAEKAYKAAEAKKAKAEKAYKAK-----AKAKKK-E 53				
		::: ::: : :::                                 :::     :::     :::				
Db	120	AEENAKQAEELKQKQAEVAAKAAADAKAAEAEAKKAAADAKKAAEAAKAAAEAKKAE 179				
QY	54	YAAAEAKYKAEAAKAAKAAEAYEA 77				
		:::                     :::     :::     :::     :::				
Db	180	VAAALKKAEAAEAAAEARKKA 203				
RESULT 2						
Q83SA1 PRELIMINARY; PRT; 413 AA.						
ID	Q83SA1					
AC	Q83SA1;					
DT	01-JUN-2003 (TReMBLrel. 24, Created)					
DT	01-JUN-2003 (TReMBLrel. 24, Last sequence update)					
DT	01-MAR-2004 (TReMBLrel. 26, Last annotation update)					
DE	Membrane spanning protein, required for outer membrane integrity.					
GN	Name=tola; OrderedLocusNames=SF0558;					
OS	Shigella flexneri.					
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;					
OC	Enterobacteriaceae; Shigella.					











```

RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01001968; EAA18109.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
SQ SEQUENCE 1701 AA; 197910 MW; 03BE665BAE45C669 CRC64;

Query Match 39.1%; Score 143; DB 2; Length 1701;
Best Local Similarity 55.3%; Pred. No. 0.012;
Matches 47; Conservative 7; Mismatches 17; Indels 14; Gaps 5;

QY 1 AKKYAKKEKA-YAKKAEKAAKAEAKYKAAKAEKKA-----EAKYKAAKAEKKA-----AEK 51
DB 1270 AKKAEKKKAAKAEKAEKKEKKAEA-AKKAEKKKAAKAEKKEKKAEEKKAEEKKAEEK 1328
QY 52 KEVAAA-----EAKYKAEKAAKAAKE 72
DB 1329 KKAEEKKAEEKKAEEKKAEEK 1353

RESULT 13
Q9WWX1 PRELIMINARY; PRT; 372 AA.
AC Q9WWX1
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE Biopolymer transport protein TolA.
DE TolA protein.
GN Name:tolA; OrderedLocusNames=PP1221;
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RA Rodriguez-Herva J.J., Ramos-Gonzalez M.I., Ramos J.;
RX MEDLINE=961981174; PubMed=8626299;
RT "The Pseudomonas putida peptidoglycan-associated outer membrane
RT lipoprotein (PAL) is involved in maintenance of the integrity of the
RT cell envelope.";
RL J. Bacteriol. 178:1699-1706(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RX MEDLINE=96422022; PubMed=8824639;
RA Rodriguez-Herva J.J., Ramos J.;
RT Characterization of an OprL null mutant of Pseudomonas putida.";
RL J. Bacteriol. 178:5836-5840(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RA Ramos-Gonzalez J.;
RT Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
RL [4]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RA Rodriguez-Herva J.J.;
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; X74218; CAB50780.1; -.
DR HSSP; P50600; ILR0.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
GO; GO:0008565; F:protein transporter activity; IEA.

```

```

DR GO; GO:0006334; P:nucleosome assembly; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR InterPro; IPR005819; Histone_H5.
DR EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
DR preliminary data.
DR InterPro; IPR006260; TonB_C.
DR Pfam; PF06519; TolA; 1.
DR PRINTS; PR00624; HISTONEH5.
DR TIGRFAMs; TIGR01352; tonB_Cterm; 1.
SQ SEQUENCE 372 AA; 40133 MW; 87F49785ECC3C0BC CRC64;

Query Match 38.9%; Score 142.5; DB 2; Length 372;
Best Local Similarity 46.9%; Pred. No. 0.0039;
Matches 46; Conservative 10; Mismatches 21; Indels 21; Gaps 4;

QY 1 AKKYAKKEKA-----YAKKAEKAAKKA-----EAKYKAAEAKKAKAKAKYA--- 44
DB 148 AKKAEKQADTAKKAEDEAKKAEKAAEAKKAAEAKKAAEDAKKAAEAKKAAED 207
QY 45 ---KAAKAEKYYAAAEAKYK--AEAKAAAEKAAEAYEA 77
DB 208 AKKKAADAEAKKAAEAKKAAADAKKKAQAKA 245

RESULT 14
Q88NI6 PRELIMINARY; PRT; 372 AA.
AC Q88NI6
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Biopolymer transport protein TolA.
GN Name:tolA; OrderedLocusNames=PP1221;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazzez A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hohseil J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AS016778; AAN66845.1; -.
DR HSSP; P50600; ILR0.
DR TIGR; PP1221; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0006334; P:nucleosome assembly; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR InterPro; IPR005819; Histone_H5.
DR InterPro; IPR010528; TolA.
DR InterPro; IPR006260; TonB_C.
DR Pfam; PF06519; TolA; 1.
DR PRINTS; PR00624; HISTONEH5.
DR TIGRFAMs; TIGR01352; tonB_Cterm; 1.
KW Complete proteome.
SQ SEQUENCE 372 AA; 40133 MW; 87F49785ECC3C0BC CRC64;

Query Match 38.9%; Score 142.5; DB 2; Length 372;
Best Local Similarity 46.9%; Pred. No. 0.0039;

```

Matches 46; Conservative 10; Mismatches 21; Indels 21; Gaps 4;

QY 1 AKKAYAKKEA-----YAKKAEKAAKKA---EAKAYKAAEAKKKKAAEAKKYA--- 44  
|||:|:|:  
Db 148 AKKAAEKQADITAKKAEDEAKKAAEEAKKAAAEADAKKAAEAEAKKKAED 207  
|||:|:|:  
QY 45 ----KAAKAEKYYAAAEAYK--AEAKAAAEAYEA 77  
|||:|:|:  
Db 208 AKKGAADAEAKKAAEAEAKKAAADAQKKAQEAARKA 245  
|||:|:|:

RESULT 15  
Q6D7F3 PRELIMINARY; PRT; 395 AA.  
AC Q6D7F3;  
DT 01-OCT-2004 (TrEMBLrel. 28, Created)  
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Tola protein.  
GN Name=tola; ORFNames=ECA1372;  
OS Erwinia carotovora subsp. atroseptica SCRI1043.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Pectobacterium.  
OX NCBI\_TaxID=218491;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SCRI1043;  
RA Bell K.S., Sebaihia M., Pritchard L., Holden M., Hyman L.J.,  
RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,  
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,  
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,  
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,  
RA Salmond G.P.C., Birch P.R.J., Barrall B.G., Parkhill J., Toth I.K.;  
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BX950851; CAG74282.1; -  
SQ SEQUENCE 395 AA; 41698 MW; 2DBIDA64E984199 CRC64;

Query Match 38.8%; Score 142; DB 2; Length 395;  
Best Local Similarity 51.1%; Pred. No. 0.0044;  
Matches 46; Conservative 9; Mismatches 17; Indels 18; Gaps 4;

QY 5 AKKAYAKKEA-----YAKKAEKAAKKA---EAKAYKAAEAKKKKAAEAKKYA--- 53  
|||:|:|:  
Db 134 AQQAEEQKQAEAAAKAAEAEQQTAAADARKKAEDEAKKQAVAAAAAKKQAEAEKE 193  
|||:|:|:  
QY 54 YAAAEAKY-----KAAEAKAAAEAYEA 77  
|||:|:|:  
Db 194 KAAEAKQKAAETAKAAEAAKAA--EAKKEA 222  
|||:|:|:

Search completed: December 14, 2004, 06:10:18  
Job time : 86.4219 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 85.1224 Seconds  
(without alignments)  
324.499 Million cell updates/sec

Title: US-10-792-311-5  
Perfect score: 366  
Sequence: 1 AKKYAKBKAVAKAEKAAK.....EAKYKAEAKAAKAEAYEA 77

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_213Sep04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	366	100.0	77	3 AAY82575	Copolymer
2	351.5	96.0	86	3 AAY82576	Copolymer
3	289	79.0	109	3 AAY82577	Copolymer
4	282.5	77.2	66	3 AAY82574	Copolymer
5	245.5	67.1	56	3 AAY82573	Copolymer
6	150.5	41.1	421	6 ABU28559	Protein e
7	148	40.4	407	6 ABU47123	Protein e
8	146.5	40.0	387	6 ABU47848	Protein e
9	146	39.9	376	6 AAY98499	Peptide #
10	144.5	39.5	100	3 AAY59044	Amino aci
11	144.5	39.5	100	3 AAY59044	Amino aci
12	144.5	39.5	100	4 AAB45852	Nucleic a
13	144.5	39.5	100	4 AAB45852	Nucleic a
14	144	39.3	428	6 ABU27824	Protein e
15	142.5	38.9	154	2 AAR06445	Recombina
16	142.5	38.9	372	6 ABU40185	Protein e
17	141.5	38.7	323	6 ABU31397	Protein e
18	141.5	38.7	347	6 ABU38313	Protein e
19	141.5	38.7	347	6 ABU18771	Pseudomon
20	141.5	38.7	407	7 ABO80835	Pseudomon
21	141.5	38.7	469	7 ABO67048	Klebsiell
22	141	38.5	372	5 ABO680418	Haemophil
23	141	38.5	372	7 ABO23507	Haemophil
24	140	38.3	106	2 AAR06446	Recombina
25	140	38.3	389	6 ABU39221	Protein e

26	138.5	37.8	357	6 ABM67869	Abm67869 Photorhab
27	134	36.6	643	2 AAR84568	Aar84568 Trypanoso
28	132	36.1	507	8 ADQ36692	Adq36692 Trypanoso
29	132	36.1	717	8 ADQ36678	Adq36678 Trypanoso
30	131.5	35.9	336	6 ABU42038	Protein e
31	131	35.8	45	3 AAY82572	Copolymer
32	130.5	35.7	388	6 ABU50266	Protein e
33	130	35.5	205	3 AAB20575	Mycobacte
34	130	35.5	472	2 AAR84569	Aar84569 Trypanoso
35	130	35.5	564	2 AAR84565	Aar84565 Trypanoso
36	129.5	35.4	214	2 AAY34055	Aay34055 M. tuberc
37	129.5	35.4	214	2 AAY57353	M. tuberc
38	129.5	35.4	214	6 ABU34623	Protein e
39	129.5	35.4	214	6 ABU36893	Protein e
40	129.5	35.4	361	7 ADF05105	Bacterial
41	128.5	35.1	214	6 ADA33882	Acinetoba
42	127	34.7	80	5 ABG71044	Tumour ne
43	125.5	34.3	468	6 ADA35034	Acinetoba
44	124.5	34.0	452	6 ABP57088	Mouse MKP
45	124	33.9	212	6 ABU33811	Protein e

ALIGNMENTS

RESULT 1  
AAY82575  
ID AAY82575 standard; peptide; 77 AA.  
XX  
AC AAY82575;  
XX  
DT 28-JUL-2000 (first entry)  
XX  
DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:5.  
XX  
KW Copolymer; molecular weight marker; TV-marker; immune disease;  
KW glatiramer acetate; autoimmune disease; antirheumatic; neuroprotective;  
KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;  
KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;  
KW inflammatory; immunosuppressive; demyelinating disease; osteoarthritis;  
KW Crohn's disease; chronic immune thrombocytopenia purpura; arthritis;  
KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;  
KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;  
KW pemphigus vulgaris; systemic lupus erythematosus.  
XX  
OS Unidentified.

XX  
FN WO200018794-A1.  
XX  
PD 06-APR-2000.  
XX  
PF 24-SEP-1999; 99WO-US022402.  
XX  
PR 25-SEP-1998; 98US-0101693P.  
XX  
PA (YEDA ) YEDA RES & DEV CO LTD.  
XX  
PI (TEVA-) TEVA PHARM USA INC.  
XX  
PI Gad A, Lis D;  
XX  
XX WPI; 2000-317499/27.  
XX  
DR Copolymer 1 related polypeptides used as molecular weight markers for  
XX glatiramer acetate and for treatment and prevention of immune diseases.  
XX  
PT Claim 10; Page 14; 72pp; English.  
XX  
PS AAY82571 to AAY82577 represent specifically claimed copolymer molecular  
XX weight TV-marker polypeptides from the present invention. The present  
XX invention describes polypeptides (I) for determining the molecular weight  
XX of a copolymer (CP), which has an identified molecular weight and an  
XX amino acid composition corresponding to the copolymer. The polypeptides











CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 387 AA;

Query Match 40.0%; Score 146.5; DB 6; Length 387;  
 Best Local Similarity 58.6%; Pred. No. 9.2e-06;  
 Matches 51; Conservative 7; Mismatches 16; Indels 13; Gaps 6;  
 QY 1 AKKYAKKEKAYAKAEKAAKAEKAYK-AAEAKKKAKAEKKA---KAAKAEKKEVA 55  
 DB 141 AKKKAPEA--AKAADAKKKAEAEVKAADAKKKAEAEAAKADAKKKAEAEAAK-A 197  
 QY 56 ABAKAYK--AEAKAA--KEAAYEA 77  
 DB 198 ABAKKAEEAEAKAAADAKKADA 224

RESULT 9  
 ABU47848  
 ID ABU47848 standard; protein; 376 AA.  
 XX  
 AC ABU47848;  
 XX  
 DT 19-JUN-2003 (first entry)  
 DE Protein encoded by Prokaryotic essential gene #33375.  
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 KW Salmomella typhi.  
 XX  
 OS WO200277183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002WO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 DR WPI; 2003-029926/02.  
 DR N-PSDB; ACA51718.  
 XX  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 PS Claim 25; SEQ ID NO 75772; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC antisense nucleic acid; (4) an antibody whose expression is inhibited by the  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 376 AA;

Query Match 39.9%; Score 146; DB 6; Length 376;  
 Best Local Similarity 54.4%; Pred. No. 9.9e-06;  
 Matches 49; Conservative 7; Mismatches 14; Indels 20; Gaps 5;  
 QY 1 AKKYAKKEKAYAKAEKAAKAEKAYK-AAEAKKKAKAEKKA---KAAKAEKKEVA 54  
 DB 158 AKKKAPEA--AKAADAKKKAEAEVKAADAKKKAEAEAAKADAKKKAEAEAAK 213  
 QY 55 ABAKAYKAEAA-----KAAKAEAA 74  
 DB 214 ABAEAKKKADAAAKAAADAKKKAEAA 243

RESULT 10  
 AAY98499  
 ID AAY98499 standard; peptide; 100 AA.  
 XX  
 AC AAY98499;  
 XX  
 DT 31-JUL-2000 (first entry)  
 DE Peptide #10 used in nucleic acid transporter system.  
 XX  
 KW Transporter system; nucleic acid delivery; gene therapy; cancer;  
 KW carcinogenesis; cardiovascular disease; infection.  
 XX  
 OS Synthetic.  
 XX  
 PN US6033884-A.  
 XX  
 PD 07-MAR-2000.  
 XX  
 PF 14-DEC-1993; 93US-00167641.  
 XX  
 PR 20-MAR-1992; 92US-00855389.  
 PR 19-MAR-1993; 93WO-US002725.  
 XX  
 PA (BAYU ) BAYLOR COLLEGE MEDICINE.

XX Gottchalk S, Sparrow J, Cristiano RJ, Woo SLC, Smith LC;  
 XX WPI; 2000-281993/24.  
 XX System for transporting nucleic acid into cells, useful e.g. in gene  
 PT therapy and for generating transgenic animals, comprises binding agent  
 PT linked to nucleic acid, surface ligand and lytic agent.  
 XX Disclosure; Col 125-128; 108pp; English.  
 XX The present invention relates to a transporter system for delivering  
 CC nucleic acid to a cell. The system comprises a nucleic acid binding  
 CC complex, consisting of a binding molecule bonded non-covalently to the  
 CC nucleic acid, and covalently to a surface ligand, and a lytic agent. The  
 CC binding molecule is spermine or a spermidine derivative. Nucleotide  
 CC sequences AAA36633-A36652 and peptide sequences AAY98456-Y98500 are used  
 CC in the construction of the transporter system of the invention. The  
 CC transporter system is used in gene therapy, particularly to deliver  
 CC nucleic acids to hepatocytes, muscle cells or bone forming cells, e.g. for  
 CC treating cardiovascular disease, cancer, and infection. The transporter  
 CC systems are also used to create transgenic animals (as models for human  
 CC carcinogenesis or disease or for drug testing). Other uses include  
 CC transforming cells to produce proteins, or transfecting cells in vitro  
 CC to study the function of the nucleic acid. The use of a surface ligand  
 CC allows specific targeting of selected cells and tissues. The lytic agent  
 CC provides for release of the nucleic acid into the cellular interior, from  
 CC endosomes, without requiring endosomal or lysosomal degradation  
 XX Sequence 100 AA;  
 SQ

Query Match 39.5%; Score 144.5; DB 3; Length 100;  
 Best Local Similarity 58.4%; Pred. No. 3.3e-06;  
 Matches 45; Conservative 6; Mismatches 23; Indels 3; Gaps 3;  
 QY 1 AKKYAK-KEKAYAKAEKAAKAEKAYKAAEAKKAKAEKAYAKA-AKAEKKEYAAAE 58  
 DB 2 AK 61  
 QY 59 AKYKAE-AKAAAEKAA 74  
 DB 62 AK 78

RESULT 11  
 AAY59044  
 ID AAY59044 standard; peptide; 100 AA.  
 AC AAY59044;  
 DT 07-MAR-2000 (first entry)  
 DE Amino acid polymer seq ID NO: 64 of US5994109.  
 XX Nucleic acid transport system; NTS; cell surface receptor; cytosol;  
 KW nuclear membrane; lysis moiety; transgenic animal; human disease;  
 KW nucleic acid delivery; cancer.  
 XX Synthetic.  
 XX Key Location/Qualifiers  
 FH Misc-difference 3..100  
 FT /notes "Lys-Ala in positions 3 to 100 may be optionally  
 FT absent"  
 FT  
 PN US5994109-A.  
 XX 30-NOV-1999.  
 PD  
 XX 03-JUN-1995; 95US-00460890.  
 PF  
 XX 20-MAR-1992; 92US-00855389.  
 PR 19-MAR-1993; 93US-0002725.  
 PR

PR 14-DEC-1993; 93US-00167641.  
 XX (BAYU ) BAYLOR COLLEGE MEDICINE.  
 XX Woo SLC, Cristiano RJ, Gottchalk S, Sparrow J, Smith LC;  
 XX WPI; 2000-038262/03.  
 DR Nucleic acid transport system, useful for creating transgenic animals for  
 XX assessing human disease such as cancer in an animal model.  
 PT Disclosure; Col 123-124; 107pp; English.  
 PS The invention relates to a nucleic acid transport system (NTS) for  
 XX delivering nucleic acid into a cell. The NTS contains but is not limited  
 CC to 5 components: (a) the nucleic acid or a macromolecule to be delivered;  
 CC (b) a moiety that recognizes and binds to a cell surface receptor or  
 CC antigen or is capable of entering a cell through cytosol; (c) a nucleic  
 CC acid or macromolecular molecule binding moiety; (d) a moiety that is  
 CC capable of moving or initiating movement through a nuclear membrane; and/  
 CC or (e) a lysis moiety that enables the transport of the entire complex  
 CC from the cell surface directly into the cytoplasm of the cell. The NTS  
 CC delivers nucleic acid into the cellular interior as well as the nucleus  
 CC of specific cells. The NTS can be used to treat disorders by targeting  
 CC specific nucleic acid accordingly. The NTS can also be used to create  
 CC transgenic animals for assessing human disease, such as cancer, in an  
 CC animal model. The NTS can be used in vitro with tissue culture cells  
 CC which allows the role of various nucleic acids to be studied by targeting  
 CC specific expression into specifically targeted tissue culture cells. The  
 CC lysis agent within the NTS avoids the problem of endosomal/lysosomal  
 CC degradation  
 XX Sequence 100 AA;  
 SQ

Query Match 39.5%; Score 144.5; DB 3; Length 100;  
 Best Local Similarity 58.4%; Pred. No. 3.3e-06;  
 Matches 45; Conservative 6; Mismatches 23; Indels 3; Gaps 3;  
 QY 1 AKKYAK-KEKAYAKAEKAAKAEKAYKAAEAKKAKAEKAYAKA-AKAEKKEYAAAE 58  
 DB 2 AK 61  
 QY 59 AKYKAE-AKAAAEKAA 74  
 DB 62 AK 78

RESULT 12  
 AAB45852  
 ID AAB45852 standard; protein; 100 AA.  
 XX AAB45852;  
 AC AAB45852;  
 DT 21-MAR-2001 (first entry)  
 DE Nucleic acid transporter system peptide ligand SEQ ID NO 64.  
 XX Nucleic acid delivery; nucleic acid transporter system; hormone; enzyme;  
 KW growth factor; clotting factor; apolipoprotein; receptor; drug; oncogene;  
 KW tumor antigen; tumor suppressor; viral antigen; parasitic antigen;  
 KW bacterial antigen.  
 XX Unidentified.  
 OS US6150168-A.  
 PN 21-NOV-2000.  
 PD  
 XX 05-JUN-1995; 95US-00460971.  
 PF  
 XX 20-MAR-1992; 92US-00855389.  
 PR 19-MAR-1993; 93US-0002725.  
 PR 14-DEC-1993; 93US-00167641.  
 PR

```

XX PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX PI Gottchalk S, Sparrow J, Cristiano RJ, Smith LC, Woo SLC;
XX DR WPI; 2001-049093/06.
XX PT Nucleic acid transporter system for delivering nucleic acid into a cell,
XX PT useful for delivering proteins and polypeptides to cells, including
XX PT growth factors, enzymes, hormones, and tumor suppressors.
XX PS Disclosure; Col 125-126; 105pp; English.
XX CC This invention describes a novel system (I) for delivering a nucleic acid
XX CC to a cell, comprising a binding complex comprising a ligand binding
XX CC molecule noncovalently bound to a nucleic acid and covalently linked to a
XX CC surface ligand, and a second binding complex comprising a second binding
XX CC molecule noncovalently bound to a nucleic acid and covalently linked to a
XX CC nuclear ligand. The complexes are simultaneously bound to the nucleic
XX CC acid. The nucleic acid transporter system can also be used in a method
XX CC for the in vivo targeting of the insertion of DNA into a cell. It can
XX CC also be used in processes for producing transformed cell lines. The
XX CC system can be used to deliver a variety of proteins and polypeptides,
XX CC such as hormones, growth factors, enzymes, clotting factors,
XX CC apolipoproteins, receptors, drugs, oncogenes, tumor antigens, tumor
XX CC suppressors, viral antigens, parasitic antigens, and bacterial antigens.
XX CC The transporter system uses lysis agents to overcome the problems of
XX CC endosomal/lysosomal degradation seen with prior art systems
XX SQ Sequence 100 AA;
XX Query Match 39.5%; Score 144.5; DB 4; Length 100;
XX Best Local Similarity 58.4%; Pred. No. 3.3e-06;
XX Matches 45; Conservative 6; Mismatches 23; Indels 3; Gaps 3;
Qy 1 AKKYAK-KEKAYAKAEKAAKAEKAYKAAEAKKKAKAEAKKYAKA-AKAEKKEYAAAE 58
Db 2 AKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAK 61
Qy 59 AKYKAEK-AKAAKAEKAA 74
Db 62 AKAKAKAKAKAKAKAKA 78
RESULT 14
ABU27824
ID ABU27824 standard; protein; 428 AA.
XX AC ABU27824;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #13351.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Enterobacter cloacae.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX PR WPI; 2003-029926/02.
XX DR N-PSDB; ACA31694.
XX PT New antisense nucleic acids, useful for identifying proteins or screening

```

```

XX PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX PI Gottchalk S, Sparrow J, Cristiano RJ, Smith LC, Woo SLC;
XX DR WPI; 2001-049093/06.
XX PT Nucleic acid transporter system for delivering nucleic acid into a cell,
XX PT useful for delivering proteins and polypeptides to cells, including
XX PT growth factors, enzymes, hormones, and tumor suppressors.
XX PS Disclosure; Col 125-126; 105pp; English.
XX CC This invention describes a novel system (I) for delivering a nucleic acid
XX CC to a cell, comprising a binding complex comprising a ligand binding
XX CC molecule noncovalently bound to a nucleic acid and covalently linked to a
XX CC surface ligand, and a second binding complex comprising a second binding
XX CC molecule noncovalently bound to a nucleic acid and covalently linked to a
XX CC nuclear ligand. The complexes are simultaneously bound to the nucleic
XX CC acid. The nucleic acid transporter system can also be used in a method
XX CC for the in vivo targeting of the insertion of DNA into a cell. It can
XX CC also be used in processes for producing transformed cell lines. The
XX CC system can be used to deliver a variety of proteins and polypeptides,
XX CC such as hormones, growth factors, enzymes, clotting factors,
XX CC apolipoproteins, receptors, drugs, oncogenes, tumor antigens, tumor
XX CC suppressors, viral antigens, parasitic antigens, and bacterial antigens.
XX CC The transporter system uses lysis agents to overcome the problems of
XX CC endosomal/lysosomal degradation seen with prior art systems
XX SQ Sequence 100 AA;
XX Query Match 39.5%; Score 144.5; DB 4; Length 100;
XX Best Local Similarity 58.4%; Pred. No. 3.3e-06;
XX Matches 45; Conservative 6; Mismatches 23; Indels 3; Gaps 3;
Qy 1 AKKYAK-KEKAYAKAEKAAKAEKAYKAAEAKKKAKAEAKKYAKA-AKAEKKEYAAAE 58
Db 2 AKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAK 61
Qy 59 AKYKAEK-AKAAKAEKAA 74
Db 62 AKAKAKAKAKAKAKAKA 78
AAU04289
ID AAU04289 standard; peptide; 100 AA.
XX AC AAU04289;
XX DT 23-OCT-2001 (first entry)
XX DE Poly-Lys-Ala used in nucleic acid transporter system.
XX KW Nucleic acid transport; cytolysis; ligand; lysis agent; spacer molecule;
XX KW gene therapy; hepatocyte; muscle; bone forming cell.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 3..100
XX FT /note= "Lys-Ala in positions 3-100 may be present or
XX FT absent"
XX PN US6177554-B1.
XX PD 23-JAN-2001.
XX PF 05-JUN-1995; 95US-00462040.
XX PR 20-MAR-1992; 92US-00855389.
XX PR 19-MAR-1993; 93WO-US0002725.
XX PR 14-DEC-1993; 93US-00167641.

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:52:22 ; Search time 191.038 Seconds  
(without alignments)  
143.965 Million cell updates/sec

Title: US-10-792-311-5  
Perfect score: 366  
Sequence: 1 AKKYAKKEKAYAKAEKAAK.....EAKYKAEAKAAKAEAYEA 77

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	366	100.0	77	9 US-09-816-989A-5	Sequence 5, Appli
2	351.5	96.0	86	9 US-09-816-989A-6	Sequence 6, Appli
3	289	79.0	109	9 US-09-816-989A-7	Sequence 7, Appli
4	282.5	77.2	66	9 US-09-816-989A-4	Sequence 4, Appli
5	245.5	67.1	56	9 US-09-816-989A-3	Sequence 3, Appli
6	150.5	41.1	421	15 US-10-282-122A-56483	Sequence 56483, A
7	148	40.4	407	15 US-10-282-122A-75047	Sequence 75047, A
8	146.5	40.0	387	15 US-10-282-122A-72645	Sequence 72645, A
9	146	39.9	376	15 US-10-282-122A-75772	Sequence 75772, A
10	144	39.3	428	15 US-10-282-122A-55748	Sequence 55748, A
11	142.5	38.9	372	15 US-10-282-122A-68109	Sequence 68109, A
12	141.5	38.7	323	15 US-10-282-122A-59321	Sequence 59321, A
13	141.5	38.7	347	14 US-10-127-032-120	Sequence 120, App

14	141.5	38.7	347	15	US-10-282-122A-66237	Sequence 66237, A
15	141	38.5	372	10	US-09-820-843A-8	Sequence 8, Appli
16	141	38.5	372	16	US-10-467-421-16	Sequence 16, Appli
17	140	38.3	389	15	US-10-282-122A-67145	Sequence 67145, A
18	132	36.1	507	16	US-10-726-692-50	Sequence 50, Appli
19	132	36.1	717	15	US-10-726-692-36	Sequence 36, Appli
20	131.5	35.9	336	15	US-10-282-122A-69962	Sequence 69962, A
21	131	35.8	45	9	US-09-816-989A-2	Sequence 2, Appli
22	130.5	35.7	388	15	US-10-282-122A-78190	Sequence 78190, A
23	129.5	35.4	214	14	US-10-229-567-27	Sequence 27, Appli
24	129.5	35.4	214	15	US-10-282-122A-62547	Sequence 62547, A
25	129.5	35.4	214	15	US-10-282-122A-64817	Sequence 64817, A
26	124.5	34.0	452	14	US-10-184-832-5	Sequence 5, Appli
27	124	33.9	212	15	US-10-282-122A-61735	Sequence 61735, A
28	123	33.6	433	15	US-10-390-472-2	Sequence 2, Appli
29	123	33.6	448	15	US-10-282-122A-45264	Sequence 45264, A
30	122.5	33.5	329	15	US-10-282-122A-67699	Sequence 67699, A
31	121.5	33.2	239	15	US-10-282-122A-60543	Sequence 60543, A
32	121	33.1	223	13	US-10-051-643-201	Sequence 201, App
33	121	33.1	223	14	US-10-205-979-52	Sequence 52, Appli
34	121	33.1	376	14	US-10-156-761-9889	Sequence 9889, Ap
35	120	32.8	240	17	US-10-739-930-9712	Sequence 9712, Ap
36	120	32.8	272	17	US-10-739-930-10710	Sequence 10710, A
37	120	32.8	568	17	US-10-435-115-197135	Sequence 197135, A
38	120	32.8	636	15	US-10-425-114-37076	Sequence 37076, A
39	120	32.8	926	16	US-10-437-963-193381	Sequence 193381, A
40	117	32.0	217	14	US-10-156-761-10221	Sequence 10221, A
41	117	32.0	279	14	US-10-181-071-7	Sequence 7, Appli
42	117	32.0	827	16	US-10-437-963-152005	Sequence 152005, A
43	116.5	31.8	885	17	US-10-425-115-198446	Sequence 198446, A
44	114	31.1	127	17	US-10-425-115-237481	Sequence 237481, A
45	112.5	30.7	146	15	US-10-296-115-1023	Sequence 1023, Ap

ALIGNMENTS

RESULT 1

US-09-816-989A-5

; Sequence 5, Application US/09816989A

; Patent No. US20020115103A1

; GENERAL INFORMATION:

; APPLICANT: Gad, Alexander

; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK

; FILE REFERENCE: 2609/60807-A-PCT-US

; CURRENT FILING DATE: 2001-03-23

; PRIOR FILING DATE: 1998-09-25

; PRIOR APPLICATION NUMBER: PCT/US99/22402

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 5

; LENGTH: 77

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide  
US-09-816-989A-5

Query Match	100.0%;	Score 366;	DB 9;	Length 77;
Best Local Similarity	100.0%;	Pred. No. 9e-25;	Mismatches 0;	Indels 0;
Matches	77;	Conservative 0;	Gaps 0;	
QY	1	AKKYAKKEKAYAKAEKAAKAEKAYKAAEAKKAKAEAKKYAKAAAEKKEYYAAAEAK	60	
Db	1	AKKYAKKEKAYAKAEKAAKAEKAYKAAEAKKAKAEAKKYAKAAAEKKEYYAAAEAK	60	
QY	61	YKAEAKAAAEKAAEAYEA 77		







```

; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55748
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55748

Query Match      39.3%; Score 144; DB 15; Length 428;
Best Local Similarity 59.0%; Pred. No. 8.3e-05;
Matches 49; Conservative 11; Mismatches 11; Indels 12; Gaps 6;

QY 1 AKYA--KKEKAYAKAEKAA-----KKAERAKYK--AAEAKKAKAEAKYKAAKAEKK 52
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 167 AKKAADAQKKAEEAEAKKAADAQKKAEEAEAKKAADAQKKAEEAEAKKA--AQEAEEK 225
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 53 EYAAAEAKYKAEAA--KAAAEKAA 74
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 226 --AAAEAKKAAAEAKKAA 246
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 11
US-10-282-122A-68109
; Sequence 68109, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55748
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55748

Query Match      39.3%; Score 144; DB 15; Length 428;
Best Local Similarity 59.0%; Pred. No. 8.3e-05;
Matches 49; Conservative 11; Mismatches 11; Indels 12; Gaps 6;

QY 1 AKYA--KKEKAYAKAEKAA-----KKAERAKYK--AAEAKKAKAEAKYKAAKAEKK 52
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 167 AKKAADAQKKAEEAEAKKAADAQKKAEEAEAKKAADAQKKAEEAEAKKA--AQEAEEK 225
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 53 EYAAAEAKYKAEAA--KAAAEKAA 74
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 226 --AAAEAKKAAAEAKKAA 246
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 12
US-10-282-122A-59321
; Sequence 59321, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68109
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-282-122A-68109

Query Match      38.9%; Score 142.5; DB 15; Length 372;
Best Local Similarity 46.9%; Pred. No. 9.7e-05;
Matches 46; Conservative 10; Mismatches 21; Indels 21; Gaps 4;

QY 1 AKYAKKEKA-----YAKKAERAKKA--AAEAKKAAEAKKAAEAKKAAEAKKAAEAK 44
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 148 AKKAAEKQADIAKKKAEDAKKAAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAAEAK 207
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 45 ---KKAKEKKEYYAAAEAKYK--AAEAKKAAEAKAEAYEA 77
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 208 AKKAAAEADAKKAAEAEAKKAAADAAQKKAQAQAARKA 245
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```

```

; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59321
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59321

Query Match          38.7%; Score 141.5; DB 15; Length 323;
Best Local Similarity 56.3%; Pred. No. 0.0001;
Matches 40; Conservative 9; Mismatches 19; Indels 3; Gaps 2;

QY 8 EKAYAKAEKAAKAEKAYK-AAEAKKKAKAKAKYAAKAEKKEKYAAAEAKYAEAA 66
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 55 EQAAKAAADAKKQAEAAAKAAAEAKKQAEAAK--AAAEAKKAEAAAKKQAEAE 112
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
QY 67 KAAAKAAAYEA 77
   |:| |:| |:| |:|
Db 113 KKAQQAANKQA 123

RESULT 13
US-10-127-032-120
; Sequence 120, Application US/10127032
; Publication No. US20030113742A1
; GENERAL INFORMATION:
; APPLICANT: Whiteley, Marvin
; APPLICANT: Bangera, M. Gita
; APPLICANT: Lory, Stephen
; APPLICANT: Greenberg, Everett Peter
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF
; FILE OF INVENTION: BIOFILM FORMATION
; FILE REFERENCE: UIZ-070CP
; CURRENT APPLICATION NUMBER: US/10/127,032
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,190
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/344,142
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 120
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-127-032-120

Query Match          38.7%; Score 141.5; DB 14; Length 347;
Best Local Similarity 44.1%; Pred. No. 0.00011;
Matches 45; Conservative 9; Mismatches 19; Indels 29; Gaps 3;

QY 1 AKKYAKKEKAYAKAEKAAKBAKA-----YKAAE-A 32
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 120 ARKAEAKAEAKKAEKAEKAEKAEKQKQADIAKKAEDAEAKKKAEDAKKKAEDA 179
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|

QY 33 KKKAKAEKKYAKAEKAEKKEKYAAAEAKYAEAAKAAKAA 74
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 180 KKKAAEEAKKKA-AAEAAKKAAVEAAKKAKEAAKAAKAA 220
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|

RESULT 14
US-10-282-122A-66237
; Sequence 66237, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
```

```

; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carz, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66237
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-66237

Query Match          38.7%; Score 141.5; DB 15; Length 347;
Best Local Similarity 44.1%; Pred. No. 0.00011;
Matches 45; Conservative 9; Mismatches 19; Indels 29; Gaps 3;

QY 1 AKKYAKKEKAYAKAEKAAKAEKA-----YKAAE-A 32
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 120 ARKAEAKAEKAEKAEKAEKAEKAEKQKQADIAKKAEDAEAKKKAEDAKKKAEDA 179
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|

QY 33 KKKAKAEKKYAKAEKAEKKEKYAAAEAKYAEAAKAAKAA 74
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 180 KKKAAEEAKKKA-AAEAAKKAAVEAAKKAKEAAKAAKAA 220
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|

RESULT 15
US-09-820-843A-8
; Sequence 8, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN
; FILE OF INVENTION: USEFUL AS ANTI-INFECTIVES
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 372
; TYPE: PRT
; ORGANISM: H. influenzae
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: outer membrane integrity protein (tolA)
```

```
; NAME/KEY: misc_feature
; OTHER INFORMATION: gi|1573353
US-09-820-843A-8

Query Match      38.5%; Score 141; DB 10; Length 372;
Best Local Similarity 54.7%; Pred. No. 0.00013;
Matches 47; Conservative 7; Mismatches 20; Indels 12; Gaps 4;

QY      1 AKKYA-----KKEKAYAKKAEKAAKK----AEAKAYKAAEAKKKAKAEAKKYAKAAKAK 51
      ||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      158 AKRLAAAKQAEEAEAKAKAAETAAQKQAEAKAKLEAEAKAKAVAEAK--AK-AEAEA 214
      ||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      52 KEYAAAEAKYKAEAKAAKAEAAEYEA 77
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      215 KAKAAAEAKAKADAEAKAAATEAKRKA 240
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: December 14, 2004, 06:59:41  
Job time : 191.038 secs

**This Page Blank (usp10)**



Result No.	Score	Query Match	Length	DB	ID	§		Description
1	366	100.0	77	4	US-09-405-743A-5		Sequence 5, Appli	
2	351.5	96.0	86	4	US-09-405-743A-6		Sequence 6, Appli	
3	289	79.0	109	4	US-09-405-743A-7		Sequence 7, Appli	
4	282.5	77.2	66	4	US-09-405-743A-4		Sequence 4, Appli	
5	245.5	67.1	56	4	US-09-405-743A-3		Sequence 3, Appli	
6	144.5	39.5	100	2	US-08-460-890A-64		Sequence 64, Appl	
7	144.5	39.5	100	3	US-08-167-641C-64		Sequence 64, Appl	
8	144.5	39.5	100	3	US-08-460-971A-64		Sequence 64, Appl	
9	144.5	39.5	100	3	US-08-462-040-64		Sequence 64, Appl	
10	141.5	38.7	407	4	US-09-252-991A-29581		Sequence 29581, A	
11	141.5	38.7	49	4	US-09-489-038A-13565		Sequence 13565, A	
12	134	36.6	643	3	US-08-216-894-8		Sequence 8, Appli	
13	134	36.6	643	3	US-09-115-746-8		Sequence 8, Appli	
14	131	35.8	45	4	US-09-405-743A-2		Sequence 2, Appli	
15	130	35.5	472	2	US-08-216-894-10		Sequence 10, Appl	
16	130	35.5	472	3	US-09-115-746-10		Sequence 10, Appl	
17	130	35.5	564	2	US-08-216-894-2		Sequence 2, Appli	
18	130	35.5	564	3	US-09-115-746-2		Sequence 2, Appli	
19	129.5	35.4	214	4	US-09-041-889-27		Sequence 27, Appl	
20	129.5	35.4	214	3	US-09-417-264-27		Sequence 27, Appl	
21	129.5	35.4	361	4	US-09-543-681A-5390		Sequence 5390, Ap	
22	128.5	35.1	214	4	US-09-328-352-5169		Sequence 5169, Ap	
23	125.5	34.3	468	4	US-08-328-352-6321		Sequence 6321, Ap	
24	123	33.6	433	1	US-08-346-849-2		Sequence 2, Appli	
25	123	33.6	433	2	US-08-293-284A-2		Sequence 2, Appli	
26	123	33.6	433	3	US-08-898-300-2		Sequence 2, Appli	
27	121	33.3	223	3	US-09-095-855-201		Sequence 201, App	



RESULT 7  
US-08-167-641C-64  
; Sequence 64, Application US/08167641C  
; Patent No. 6033884  
; GENERAL INFORMATION:  
; APPLICANT: Woo, Savio L.C.  
; APPLICANT: Smith, Louis C.  
; APPLICANT: Cristiano, Richard J.  
; APPLICANT: Gottchalk, Stephen  
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND  
; TITLE OF INVENTOR: METHODS OF USE  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSeq for Windows 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/167,641C  
; FILING DATE: December 14, 1993

RESULT 8  
 US-08-460-971A-64  
 ; Sequence 64, Application US/08460971A  
 ; Patent No. 6150168  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Woo, Savio L.C.  
 ; APPLICANT: Smith, Louis C.  
 ; APPLICANT: Cristiano, Richard J.  
 ; APPLICANT: Gottchalk, Stephen  
 ; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND  
 ; TITLE OF INVENTION: METHODS OF USE  
 ; NUMBER OF SEQUENCES: 65  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Lyon & Lyon  
 ; STREET: 633 West Fifth Street  
 ; STREET: Suite 4700  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: U.S.A.  
 ; ZIP: 90071-2066  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 ; MEDIUM TYPE: storage  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: IBM P.C. DOS 5.0  
 ; SOFTWARE: FastSEQ for Windows 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/460,971A  
 ; FILING DATE: June 5, 1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/167,641  
 ; FILING DATE: December 14, 1993  
 ; APPLICATION NUMBER: 07/855,389



```

: APPLICANT: Gary Breton et. al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
: TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 2709.2004001
: CURRENT APPLICATION NUMBER: US/09/489,039A
: CURRENT FILING DATE: 2000-01-27
: PRIOR APPLICATION NUMBER: US 60/117,747
: PRIOR FILING DATE: 1999-01-29
: NUMBER OF SEQ ID NOS: 14342
: SEQ ID NO 13565
: LENGTH: 469
: TYPE: PRT
: ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13565

```

Query Match	38.7%	Score 141.5;	DB 4;	Length 469;
Best Local Similarity Matches	56.3%;	Pred. No. 5.1e-06;		
Matches	40; Conservative	9; Mismatches	19; Indels	3; Gaps
Qy	8 EKAYAKAEKAAKGAEKAKYK-AAEAKKKAKAEEAKCYAKAAKAKEKYEAARAKYAEEA	66		
Db	205 EQAAAKAADAQKQAEEAAAANKAAAEAKKQAEEAAK--AAAEAQKAAEEAAAANKQAQEAE	262		
Qy	67 KAAAKEAAEAYA	77		
Db	263 KKAOEAAKOA	273		

RESULT 12  
US-08-216-894-8  
; Sequence 8, Application US/08216894  
; Patent No. 5876734  
; GENERAL INFORMATION:  
; APPLICANT: Kirchhoff, Louis V.  
; APPLICANT: Otsu, Keiko  
; TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION  
; TITLE OF INVENTION: WITH TRYPAOSOMA CRUZI  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington, D.C.  
; COUNTRY: USA

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/216.894
FILING DATE: 24-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 85326/102/DRLO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 643 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-216-894-8

```

Query Match 36.6%; Score 134; DB 2; Length 643;  
Best Local Similarity 52.0%; Pred. No. 3.6e-05;  
Matches 39; Conservative 9; Mismatches 23; Indels 4; Gaps 2;  
Qv 1 AKKYAKKEKAYAKAEAKKAEAKAYKAAEAKKKAKAEAKKYAKAAK-ABKKEVAAAEAA 59

[illegible]

```

Query Match      36.6%; Score 134; DB 3; Length 643;
Best Local Similarity 52.0%; Pred. No. 3.6e-05;
Matches 39; Conservative 9; Mismatches 23; Indels 4; Gaps 2;

Qy      1 AKKYAKKEKAYAKAEKAAKAAKAYAKAAEAKKKAKAAEKVAKAAK-AEKKEYAAAAA 59
      ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: |||
Db      490 ATKVAEAEK---QKAAEATTKVAEAEKQKAAEATTKVAEAEKQKAAEATTKVAEAEKQKAAEA 546

Qy      60 KYKAEAAKAAKAAAEAA 74
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      547 TKVAEAEKQKAAEAA 561

RESULT 14
US-09-405-743A-2
; Sequence 2, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co., Ltd.
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
; FILE REFERENCE: 60807-A
; CURRENT APPLICATION NUMBER: US/09/405,743A

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 19.7764 Seconds  
(without alignments)  
418.411 Million cell updates/sec

Title: US-10-792-311-6  
Perfect score: 409  
Sequence: 1 AKKYAKKEKAYAKAEKAAK.....KKYAKAEAKAAKAEAYEA 86

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	164	40.1	421	2 JV0057	tola protein - Esc
2	157.5	38.5	376	2 AG0592	tola protein [impo
3	156.5	38.3	1701	2 T09127	probable erythrocy
4	135	37.9	394	2 F90725	membrane spanning
5	155	37.9	394	2 G85576	membrane spanning
6	154	37.7	347	2 E83525	Tola protein PA097
7	151	36.9	210	2 A25550	histone H1 - sea u
8	149	36.4	231	2 S59589	histone H1 - Chlam
9	147.5	36.1	388	2 AC0138	Tola colicin impor
10	147	35.9	214	2 G70673	probable hupB - My
11	144	35.2	211	2 A28100	histone H1-beta, e
12	142.5	34.8	344	2 S34153	met101-1 protein -
13	142	34.7	372	2 G4064	outer membrane int
14	140.5	34.4	328	2 A44993	cytosolic repetiti
15	140.5	34.4	1390	2 S51364	sperm tail-specifi
16	138	33.7	217	2 A26721	histone H1-gamma,
17	137.5	33.6	1128	2 T30296	R27-2 protein - Tr
18	135	33.0	219	2 E60110	repetitive protein
19	134	32.8	206	2 S09388	histone H1 - sea u
20	133.5	32.6	384	2 B43592	outer membrane pro
21	132.5	32.6	433	2 S25194	zuotin - yeast (Sa
22	132.5	32.4	206	1 HSTR1R	histone H1 - rainb
23	132	32.3	248	1 HSUR1P	histone H1, gonada
24	131	32.0	208	2 T23778	histone H1.1 - Cae
25	130.5	31.9	220	2 A28456	histone H1.10 - ch
26	130.5	31.9	243	2 AE1689	hypothetical prote
27	129.5	31.7	924	2 T06636	hypothetical prote
28	129.5	31.7	1403	2 T11583	probable translati
29	129	31.5	218	2 A23055	histone H1.01 - ch

ALIGNMENTS

RESULT 1

JV0057  
tola protein - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004  
C:Accession: JV0057; B64810  
R:Levengood, S.K.; Webster, R.E.  
J. Bacteriol. 171, 6600-6609, 1989  
A:Title: Nucleotide sequences of the tola and tolB genes and localization of their products  
A:Reference number: JV0057; MUID:90078104; PMID:2687247  
A:Accession: JV0057  
A:Molecule type: DNA  
A:Residues: 1-421 <LEV>  
A:Cross-references: UNIPROT:PI9934; GB:M28232; NID:G148018; PIDN:AAA24683.1; PID:G148019  
A:Experimental source: strain JM105  
A>Note: The authors translated the initiation codon GTG for residue 1 as Val  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: B64810  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-421 <BLAT>  
A:Cross-references: GB:AE000177; GB:U00096; NID:G1786955; PIDN:AACT3833.1; PID:G1786960;  
A:Experimental source: strain K-12, substrain MG1655  
C:Comment: tola and tolB proteins are necessary for colicins E2, E3, A, and K to reach the  
C:Genetics:  
A:Gene: tola  
A:Map position: 17 min  
A:Start codon: GTG  
A:Keywords: nucleotide binding; P-loop; transmembrane protein  
F14-34/Domain: transmembrane #status predicted <MSS>  
F78-301/Domain: helical #status predicted <HSG>  
F:355-362/Region: nucleotide-binding motif A (P-loop)

Query Match 40.1%; Score 164; DB 2; Length 421;  
Best Local Similarity 59.3%; Pred. No. 5.3e-05;  
Matches 51; Conservative 7; Mismatches 18; Indels 10; Gaps 4;  
QY 7 KEKAYAKAEKAAKAEKAYAKAEAK-KKAEAKKAYAKAA-----KAEKKYAAAAA 59  
Db 148 KAEADAKAEAEAKKAAADAKKKAEAEAKAAAEAKKAEAEAAALKKAEAEAAAAA 207  
QY 60 KYKA--EAKKAYKAEAAKAAKEAA 83  
Db 208 RKKAATEAAEKA-KAEAEKKAEEKA 232

RESULT 2

AG0592

[illegible]



A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-347 <STO>  
A:Cross-references: UNIPROT:P50600; GB:AE004530; GB:AE004685; PIDN:AAG0436  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: tola; PA0971

Query Match 37.7%; Score 154; DB 2; Length 347;  
Best Local Similarity 47.6%; Pred. No. 0.00023;  
Matches 49; Conservative 11; Mismatches 21; Indels 22; Gaps 4;

QY 1 AKYAKKEKAYAKKAEAKKAEAKYKAA-----EAKKAKAEAKYKAA----44  
DB 120 ARKAEQAQKAAEAKKAADEAKKAAEAKQKQADIAKGAEDAEAKKAAEDAKKAAEDA 179  
QY 45 --KAABEKKEVAAAEAKYKA--EAAKAYKAEAAKAAAEKAA 83  
DB 180 KKAABEAKKGAANEAKKAAVEAAK--KAAAAAAAAAKAA 220

RESULT 7  
A25550  
histone H1 - sea urchin (Lytechinus pictus)  
C:Species: Lytechinus pictus (painted urchin)  
C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 09-Jul-2004  
C:Accession: A25550  
R:Knowles, J.A.; Childs, G.J.  
Nucleic Acids Res. 14, 8121-8133, 1986  
A:Title: Comparison of the late H1 histone genes of the sea urchins Lytechinus pictus and  
A:Reference number: A25550; MUID:87040778; PMID:3022245  
A:Accession: A25550  
A:Molecule type: DNA  
A:Residues: 1-210 <KNO>  
A:Cross-references: UNIPROT:P06144; GB:X04488; NID:g9616; PIDN:CAA28177.1; PID:g9617  
A:Superfamily: histone H1  
C:Keywords: Chromosomal protein; DNA binding; nucleosome; nucleus

Query Match 36.9%; Score 151; DB 2; Length 210;  
Best Local Similarity 51.7%; Pred. No. 0.00026;  
Matches 46; Conservative 10; Mismatches 25; Indels 8; Gaps 3;

QY 2 KKYAKKEKAYAKAEKAAKAEKAYKAAEAK-KKAKAEAKYKAAKAEKKEKAYAAEAK 60  
DB 96 KTEAQKARAANKAKLAKEKGEKKAATKARKEKLAAKKAAKKA-AKTKVKPAAKAK 154  
QY 61 YKAE-----AAKKYKAEAAKAAAEKAA 83  
DB 155 KPAAKAAKPAKAAKAAKPAKPAKAA 183

RESULT 8  
S59589  
histone H1 - Chlamydomonas reinhardtii  
C:Species: Chlamydomonas reinhardtii  
C:Date: 10-Apr-1996 #sequence\_revision 19-Apr-1996 #text\_change 09-Jul-2004  
C:Accession: S59589; S62122  
R:Fabry, S.; Mueller, K.; Lindauer, A.; Park, P.B.; Cornelius, T.; Schmitt, R.  
Curr. Genet. 28, 333-345, 1995  
A:Title: The organization structure and regulatory elements of Chlamydomonas histone gene  
A:Reference number: S59581; MUID:96120862; PMID:8590479  
A:Accession: S59589  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-231 <FAB>  
A:Cross-references: UNIPROT:Q39576; EMBL:U16726  
R:Fabry, S.; Ford, C.; Lindauer, A.; Mueller, K.; Cornelius, T.; Schmitt, R.  
submitted to the EMBL Data Library, October 1994  
A:Description: The organization, structure and controlling elements of Chlamydomonas histone  
A:Reference number: S62122  
A:Accession: S62122  
A:Molecule type: DNA

A:Residues: 1-173, 'P', 174-231 <PAW>  
A:Cross-references: EMBL:U16726; NID:g571479; PIDN:AAA98452.1; PID:g571480  
C:Genetics:  
A:Introns: 62/3; 101/3  
C:Superfamily: histone H1  
C:Keywords: Chromosomal protein; DNA binding; nucleosome

Query Match 36.4%; Score 149; DB 2; Length 231;  
Best Local Similarity 39.5%; Pred. No. 0.00038;  
Matches 47; Conservative 7; Mismatches 29; Indels 36; Gaps 2;

QY 1 AKKYAKKEKAYAKKAEAKKAAEA-----KAVKAAEAKKKAKAEAKKYKAAKAAK 51  
DB 103 AKAAATPKAAAPKKEGAVKTKAPKKEGKPKSAKKAKEKPKKEGKKAAPKAAK 162  
QY 52 KEYAA-----AEAKYKAEAAKAYKAEAAKAAAEKAA 83  
DB 163 KPKAAKPAKTKTKAAAKPKAEKPKAAAKPKAEKKPKAAKPKAEKKAAKAAKPAAKKSA 221

RESULT 9  
AC0138  
TolA colicin import membrane protein [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AC0138  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, E.  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AC0138  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-388 <KUR>  
A:Cross-references: UNIPROT:Q8ZG22; GB:AL590842; PIDN:CAC89966.1; PID:g15979190; GSPDB:G1  
C:Genetics:  
A:Gene: tola

Query Match 36.1%; Score 147.5; DB 2; Length 388;  
Best Local Similarity 45.2%; Pred. No. 0.00072;  
Matches 47; Conservative 9; Mismatches 21; Indels 27; Gaps 3;

QY 7 KEKAYAKKAEAKKAEAKAYK----AAEAKKKAKAEAKKYKAAKAEKKE-----53  
DB 155 XAKERQKQAEATAAQAKAEADKIVKQAQAEAKKAEAEAKKEAAVAAAKKQADADAKAV 214  
QY 54 -----YAAAEAKYKAEAAK-AYKAEAAKAAAEKAA 83  
DB 215 EVAEAKAAADAEEKKAAADAEEKKAAAKKVAATAAEAKKAAAEKAA 258

RESULT 10  
G70673  
probable hupB - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: G70673  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;  
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Sgares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: G70673  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-214 <COL>  
A:Cross-references: UNIPROT:P95109; GB:283018; GB:AL123456; NID:g3261671; PIDN:CAB05427.1  
A:Experimental source: strain H37RV  
C:Genetics:



Search completed: December 14, 2004, 05:52:15  
Job time : 20.7764 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 96.5232 Seconds  
(without alignments)  
512.646 Million cell updates/sec

**Title:** US-10-792-311-6

Perfect score: 409  
Sequence: 1 AKTYAKKEKAYAKKAEKAAK.....KKAYKAEAAKAAAKEAAVEA 86

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 s

```
Database :      UniProt_02:*
            1: uniprot_sprot:*
            2: uniprot_trembl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	*				
1	165.5	40.5	40.5	407	2	Q8ZQT6	Q8zqt6 salmonella
2	165	40.3	40.3	1701	2	Q7RC08	Q7rc08 plasmodium
3	164	40.1	40.1	421	1	TOLA ECOLI	P19334 escherichia
4	163	39.9	39.9	372	2	Q9WXX1	Q9wwx1 pseudomonas
5	163	39.9	39.9	372	2	Q88N16	Q88n16 pseudomonas
6	162	39.6	41.3	413	2	Q7C2Q4	Q7c2q4 shigella fl
7	162	39.6	41.3	413	2	Q83SA1	Q83sa1 shigella fl
8	160.5	39.2	39.5	395	2	Q6D7F3	Q6d7f3 erwinia car
9	160	39.1	421	421	2	Q8FYT1	Q8fyt1 escherichia
10	157.5	38.5	38.5	376	2	Q8Z8C1	Q8z8c1 salmonella
11	156.5	38.3	38.3	1652	2	Q7KEZ1	Q7kpez1 plasmodium
12	156.5	38.3	38.3	1652	2	AAQ73456	AAq73456 plasmodium
13	156.5	38.3	38.3	1680	2	Q7KPY9	Q7kpy9 plasmodium
14	156.5	38.3	38.3	1680	2	AAQ73455	AAq73455 plasmodium
15	156.5	38.3	38.3	1680	2	AAQ73457	AAq73457 plasmodium
16	156.5	38.3	38.3	1701	2	O61164	O61164 plasmodium
17	155	37.9	37.9	379	2	Q7WFN5	Q7wfn5 bordetella
18	155	37.9	394	394	2	Q7AG18	Q7ag18 escherichia
19	155	37.9	394	394	2	Q8X965	Q8x965 escherichia
20	154	37.7	347	347	1	TOLA PSEAE	P50600 pseudomonas
21	154	37.7	373	373	2	Q7W477	Q7w477 bordetella
22	154	37.7	395	395	2	Q937K4	Q937k4 erwinia chr
23	153	37.4	401	401	2	Q74W64	Q74w64 yersinia pe
24	153	37.4	401	401	2	AA61283	AA61283 yersinia
25	152.5	37.3	713	713	2	Q6CDX0	Q6cdx0 yarrowia li
26	151	36.9	210	210	1	H1 LVPTI	P06144 lytechinus
27	151	36.9	1341	1341	2	Q98KG7	Q98kg7 rhizobium l
28	150.5	36.8	198	2	Q6MIU4	Q6mtu4 bdellovibri	
29	150.5	36.8	198	2	CAE80819	CAe80819 bdellovib	
30	148.5	36.3	232	2	Q39576	Q39576 chlamydomo	
31	148	36.2	389	2	Q9CM70	Q9cm70 pasteurella	

RESULT 1  
08Z076

Q8ZQ16 ID 08ZOT6 PRELIMINARY: PRT: 407 AA.

QZQTS;  
AC 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Tol protein, membrane spanning protein.  
GN Name=ColA; OrderedLocustNames=STM0747;  
OS *Salmonella typhimurium*.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
NCBI\_TaxID=602;  
FN [1]

## ALIGNMENTS

RP SEQUENCE FROM N.A.

RC	STRAIN=LT2;
RX	MCLELINE=21534948; PubMed=11677609;
RA	McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA	Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA	Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA	Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA	Waterston R., Wilson R.K.;
RT	"Complete genome sequence of <i>Salmonella enterica</i> serovar Typhimurium
RT	LT2.";
RL	Nature 413:852-856(2001).
DR	EWBL; AE008730; AAL19691.1; -.
DR	HSP; P19934; 1TOL.
DR	InterPro; IPR010528; TOLa.
DR	Pfam; PF06519; TOLa; 1.
KW	Complete proteome.
SO	SEQUENCE 407 AA; 41865 MW; 25343522116602D75 CRC64;

Query Match 40.5%: Score 165.5: DB 2: Length 407:

Best Local Similarity 51.9%; Pred. NO. 0.00018;  
Matches 55; Conservative 14; Mismatches 16; Indels 21; Gaps 6;

QY 1 AKTYAKTKENYAKAEKAA-----KKAERAKYK-AABAKTKAKAEAKKYA-----KAARAEK 51

Dh 123 AAKLAOOOOOAEAAKAAADAKKKAEEAAKAAADAKKKAEEAAVKAAADAKKKAEEAA 182

52 KEYAAAEAKYKAEAE-----AKKAYKAEAAKAAA--KEAAVEA 86

52 KE1AAEAKENAAE - - - - - ANGINAAEAGAG - - - - - KAAE1EA 86  
: |||: || ||||  
183 AK-AAADAKKKAEEAAKAAAEAKKKAAEAKAAEAKKKAAE 227

## RESULT 2

Q7RC08	PRELIMINARY;	PRT; 1701 AA.
ID	Q7RC08	
AC	Q7RC08;	
DT	01-MAR-2004	(TREMblrel. 26, Created)
DT	01-MAR-2004	(TREMblrel. 26, Last sequence update)
DT	01-MAR-2004	(TREMblrel. 26, Last annotation update)
DE	Erythrocyte binding protein.	
GN	Name=Py0597;	

```

OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865;
RA Carlton J.M., Anguolli S.V., Suh B.B., Kooij T.W., Perteau M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selegut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabli A., Cummings L.M.,
RA Florns L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -I- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
DR EMBL; AABL01001968; EAA18109.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duff binding.
DR Pfam; PF05424; Duff binding; 1.
SQ SEQUENCE 1701 AA; 197910 MW; 03BE665BAE45C669 CRC64;

Query Match 40.3%; Score 165; DB 2; Length 1701;
Best Local Similarity 56.1%; Pred. No. 0.00064;
Matches 55; Conservative 18; Mismatches 18; Indels 18; Gaps 6;

QY 1 AKKYAKKEKA-YAKKAEKAKKAEKAYKAAKAKKAKA-----EAKKYAKAAK---AEK 51
||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1270 AKKAEKKKAAKAAKAEKKKAAE-AKAEKKKAAKAAKAEKKKAAKAAKAAKAEK 1328

QY 52 KEVAAA-----EAKYKAEAKKA-----YKAEAAKAAAKE 81
||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1329 KKAEEAKKAEKKKAAKAAKAEKKKAAKAAKAAKAAKEE 1366

RESULT 3
TOLA_ECOLI
ID TOLA_ECOLI STANDARD; PRT; 421 AA.
AC P19934;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Tola protein.
GN Name:tola; Synonyms=cim, excC, lky; OrderedLocusNames=b0739;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / JM105;
RX MEDLINE=90078104; PubMed=2687247;
RA Levensgood S.K., Webster R.E.;
RT "Nucleotide sequences of the tola and tolB genes and localization of
RT their products components of a multistep translocation system in
RT Escherichia coli."
RL J. Bacteriol. 171:6600-6609(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

```

```

RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-Kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map."
RL DNA Res. 3:137-155(1996).
RN [4]
RP DOMAINS.
RX MEDLINE=91296736; PubMed=2068069;
RA Levensgood S.K., Beyer W.F. Jr., Webster R.E.;
RT "Tola: a membrane protein involved in colicin uptake contains an
RT extended helical region."
RL Proc. Natl. Acad. Sci. U.S.A. 88:5939-5943(1991).
RN [5]
RP INTERACTION WITH PORINS.
RX MEDLINE=97133271; PubMed=8978668;
RA Derouiche R., Gavioli M., Benedetti H., Prilipov A., Lazdunski C.,
RA Lloubes R.;
RT "Tola central domain interacts with Escherichia coli porins."
RL EMBO J. 15:6408-6415(1996).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 298-421.
RX MEDLINE=99332679; PubMed=10404600;
RA Lubkowski J., Hennecke F., Pluckthun A., Wlodawer A.;
RT "Filamentous phage infection: crystal structure of g3p in complex with
RT its coreceptor, the C-terminal domain of Tola."
RL Structure 7:711-722(1999).
CC -I- FUNCTION: Involved in the tonB-independent uptake of group A
CC colicins (colicins A, E1, E2, E3, and K). Necessary for the
CC colicins to reach their respective targets after initial binding
CC to the bacteriophage DNA.
CC bacteriophage DNA.
CC -I- SUBUNIT: Interacts, via domain II, with porins ompC, phoE and
CC lamb.
CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M28232; AAA24683.1; -.
CC EMBL; AE000177; AAC73833.1; -.
CC EMBL; D90713; BAA35405.1; -.
CC FIR; JY0057; JY0057.
CC PDB; ITOL; X-ray; A=--
CC EchoBASE; EB1000; -.
CC EcoGene; EG11007; tola.
CC InterPro; IPR010528; Tola.
CC Pfam; PF06519; Tola; 1.
KM 3D-structure; Bacteriocin transport; Complete proteome;
KW Inner membrane; Protein transport; Repeat; Transmembrane; Transport.
FT DOMAIN 1 13 Cytoplasmic (Potential).
FT TRANSMEM 14 34 Potential.
FT DOMAIN 35 421 Periplasmic (Potential).
FT DOMAIN 48 310 DOMAIN II (ALPHA-HELICAL).
FT DOMAIN 311 421 DOMAIN III (FUNCTIONAL).
FT DOMAIN 224 292 13 tandem repeats of [EDA]-K(1,2)-A(2,4).
FT REPEAT 224 229 1.
FT REPEAT 230 234 2.

```

```

FT REPEAT 235 240 3.
FT REPEAT 241 245 4.
FT REPEAT 246 250 5.
FT REPEAT 251 255 6.
FT REPEAT 256 260 7.
FT REPEAT 261 266 8.
FT REPEAT 267 271 9.
FT REPEAT 272 277 10.
FT REPEAT 278 282 11.
FT REPEAT 283 287 12.
FT REPEAT 288 292 13.
FT DISULFID 363 398
FT HELIX 335 349
FT TURN 350 351
FT TURN 353 354
FT TURN 355 358
FT TURN 359 360
FT TURN 363 369
FT TURN 371 372
FT STRAND 375 383
FT HELIX 385 397
FT HELIX 406 412
FT TURN 413 414
FT STRAND 416 421
SQ SEQUENCE 421 AA; 43156 MW; 8B2F52B4B97C655E CRC64;

Query Match 40.1%; Score 164; DB 1; Length 421;
Best Local Similarity 59.3%; Pred. No. 0.00024;
Matches 51; Conservative 18; Mismatches 18; Indels 10; Gaps 4;

QY 7 KEKAYAKKAEKAAKAYKAAEAK-KKAAEAKKAYKAA-----KAEKKEVAAAEAA 59
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 148 KAEADAKAEAEAKKAAADAKKKAEEAKAAAEAAQKKAEEAAALKKKAEEAAAEAAEA 207
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 60 KYKA--EAAKAYKAAEAAKAAEAA 83
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 208 RKAAATEAAEKA-KAEAEKAAAEKA 232
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
Q9WWX1 PRELIMINARY; PRT; 372 AA.
AC Q9WWX1
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Tola protein.
CN Name=tola;
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RX MEDLINE=96198174; PubMed=8626299;
RA Rodriguez-Herva J.J., Ramos-Gonzalez M.I., Ramos J.;
RT "The Pseudomonas putida peptidoglycan-associated outer membrane
lipoprotein (PAL) is involved in maintenance of the integrity of the
cell envelope."
RL J. Bacteriol. 178:1699-1706 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RX MEDLINE=96422022; PubMed=8824639;
RA Rodriguez-Herva J.J., Ramos J.;
RT "Characterization of an OprL null mutant of Pseudomonas putida."
RL J. Bacteriol. 178:5836-5840 (1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RA Ramos-Gonzalez I.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.

```

```

RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RA Rodriguez-Herva J.J.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; X74218; CAB50780.1; -.
DR HSSP; P50600; 1LR0.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0006334; F:nucleosome assembly; IEA.
DR GO; GO:0015031; F:protein transport; IEA.
DR InterPro; IPR005819; Histone_H5.
DR InterPro; IPR010528; TOLA.
DR InterPro; IPR006260; TonB_C.
DR Pfam; PF06519; TOLA; 1.
DR PRINTS; PRO0624; HISTONEH5.
DR TIGRFAMs; TIGR01352; tonB_Cterm; 1.
SQ SEQUENCE 372 AA; 40133 MW; 87F49785ECC3C0BC CRC64;

Query Match 39.9%; Score 163; DB 2; Length 372;
Best Local Similarity 53.1%; Pred. No. 0.00025;
Matches 52; Conservative 12; Mismatches 22; Indels 12; Gaps 5;

QY 1 AKKYAKKEKA--YAKKAE-KAAKAEAKAYKAA--EAKKKAEEAKKYA-----KAAKA 49
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 148 AKKAAEKQADIAKKKAEDEAKKAAEEAKKAAEEAKKAAEDAKKAAEEAKKKAED 207
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 50 EKKEVAAAEAKYK-AEAAKAYKAAEAAKAAEAAEAAEAA 86
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 208 AKKAAAEADAKKAAEEAKKAAADAKKKAQEAARKA 245
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
Q88NI6 PRELIMINARY; PRT; 372 AA.
AC Q88NI6
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Biopolymer transport protein Tola.
GN Name=tola; OrderedLocusNames=PP1221;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Pouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazzes A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808 (2002).
DR EMBL; AE016778; AAN66845.1; -.
DR HSSP; P50600; 1LR0.
DR TIGR; PP1221; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.

```





[illegible]

```
QY 49 AEKKEYAA--AEAKYKAEAAKKA-----YKAEAAKAAAKE 81
Db 1263 RKKAEEERKKAEEAVKKAEEAKKAEAAKKAEEKKKAEAAKKAEE 1301

RESULT 12
AAQ73456 PRELIMINARY; PRT; 1652 AA.
ID AAQ73456;
AC AAQ73456;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DR 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Erythrocyte binding protein 3.
GN MAEBL.
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium;
OC Plasmodium yoelii.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YM;
RX MEDLINE=98115903; PubMed=9448314;
RA Kappe S.H., Noe A.R., Fraser T.S., Blair P.L., Adams J.H.;
RT "A family of chimeric erythrocyte binding proteins of malaria
parasites.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:1230-1235(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=YM;
RA Kappe S.H.I., Noe A.R., Fraser T.S., Blair P.L., Adams J.H.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=YM;
RA Adams J.H.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF031886; AAQ73456.1; -.
SQ SEQUENCE 1652 AA; 193757 MW; 8567E6E558B3F75C CRC64;

Query Match 38.3%; Score 156.5; DB 2; Length 1652;
Best Local Similarity 53.5%; Pred. No. 0.0025;
Matches 53; Conservative 9; Mismatches 18; Indels 19; Gaps 6;

QY 1 AKKYAKKEXAYA---KKAEEAAKKAEEAKYKAEAAKKA-----EAKKYAKAAK--- 48
Db 1204 AAKKAEEERKKAEEAVKKAEEAKKAEAAKKAEEKKKAEAAKKAEE 1262

QY 49 AEKKEYAA--AEAKYKAEAAKKA-----YKAEAAKAAAKE 81
Db 1263 RKKAEEERKKAEEAVKKAEEAKKAEAAKKAEEKKKAEAAKKAEE 1301

RESULT 14
AAQ73455 PRELIMINARY; PRT; 1680 AA.
ID AAQ73455;
AC AAQ73455;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DR 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Erythrocyte binding protein 2.
GN MAEBL.
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium;
OC Plasmodium yoelii.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YM;
RX MEDLINE=98115903; PubMed=9448314;
RA Kappe S.H., Noe A.R., Fraser T.S., Blair P.L., Adams J.H.;
RT "A family of chimeric erythrocyte binding proteins of malaria
parasites.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:1230-1235(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=YM;
RA Kappe S.H.I., Noe A.R., Fraser T.S., Blair P.L., Adams J.H.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=YM;
RA Adams J.H.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF031886; AAQ73455.1; -.
SQ SEQUENCE 1680 AA; 197169 MW; 9B955A060B1C67A9 CRC64;

Query Match 38.3%; Score 156.5; DB 2; Length 1680;
Best Local Similarity 53.5%; Pred. No. 0.0026;
Matches 53; Conservative 9; Mismatches 18; Indels 19; Gaps 6;

QY 1 AKKYAKKEXAYA---KKAEEAAKKAEEAKYKAEAAKKA-----EAKKYAKAAK--- 48
Db 1204 AAKKAEEERKKAEEAVKKAEEAKKAEAAKKAEEKKKAEAAKKAEE 1262

QY 49 AEKKEYAA--AEAKYKAEAAKKA-----YKAEAAKAAAKE 81
Db 1263 RKKAEEERKKAEEAVKKAEEAKKAEAAKKAEEKKKAEAAKKAEE 1301

RESULT 13
Q7KPY9 PRELIMINARY; PRT; 1680 AA.
ID Q7KPY9;
AC Q7KPY9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DR 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Erythrocyte binding protein 4 (Erythrocyte binding protein 2).
GN Namesmaeb1.
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YM;
RX MEDLINE=98115903; PubMed=9448314;
RA Kappe S.H., Noe A.R., Fraser T.S., Blair P.L., Adams J.H.;
RT "A family of chimeric erythrocyte binding proteins of malaria
parasites.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:1230-1235(1998).
```

Search completed: December 14, 2004, 06:10:19  
Job time : 97.5232 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 95.0717 Seconds  
(without alignments)  
324.499 Million cell updates/sec

Title: US-10-792-311-6  
Perfect score: 409  
Sequence: 1 AKYAKKEKAYAKAEKAAK.....KKAYKAEKAAKAEAYEA 86

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	409	100.0	86	3	AAY82576 Copolymer
2	351.5	85.9	77	3	AAY82575
3	312.5	76.4	109	3	AAY82577 Copolymer
4	268	65.5	66	3	AAY82574 Copolymer
5	233	57.0	56	3	AAY82573 Copolymer
6	165.5	40.5	407	6	ABU47123 Protein e
7	184	40.1	387	6	ABU44721 Protein e
8	164	40.1	421	6	ABU28559 Protein e
9	163	39.9	154	2	AAR06445 Recombina
10	163	39.9	372	6	ABU40185 Protein e
11	161	39.4	428	6	ABU27824 Protein e
12	157.5	38.5	323	6	ABU31397 Protein e
13	157.5	38.5	376	6	ABU47848 Protein e
14	157.5	38.5	469	7	ABO67048 Klebsiell
15	157	38.4	106	2	AAR06446 Recombina
16	156	38.1	100	3	AAY98499 Peptide #
17	156	38.1	100	3	AAY59044 Amino aci
18	156	38.1	100	4	AAB45852 Nucleic a
19	156	38.1	100	4	AUU04289 Poly-Lys-
20	154	37.7	347	6	ABU38313 Protein e
21	154	37.7	347	6	ABJ18771 Pseudomon
22	154	37.7	407	7	ABO80835 Pseudomon
23	148	36.2	389	6	ABU39221 Protein e
24	147.5	36.1	388	6	ABU50266 Protein e
25	147	35.9	214	2	AAY34055 M. tuberc

26	147	35.9	214	2	AAY57353
27	147	35.9	214	6	ABU34623
28	147	35.9	214	6	ABU36893
29	147	35.9	357	6	ABM67869
30	146.5	35.8	205	3	AAB20575
31	142	34.7	372	5	ABG80418
32	142	34.7	372	7	ABO23507
33	141.5	34.6	336	6	ABU42038
34	141.5	34.6	361	7	ADF05105
35	140.5	34.4	214	6	ADA33882
36	140	34.2	468	6	ADA35034
37	138.5	33.9	80	5	ABG71044
38	137.5	33.6	448	6	ABU17340
39	137.5	33.6	472	2	AAR84569
40	137.5	33.6	507	8	ADQ36692
41	137.5	33.6	564	2	AAR84565
42	137.5	33.6	643	2	AAR84568
43	137.5	33.6	717	8	ADQ36678
44	135	33.0	140	2	AAR25206
45	133.5	32.6	433	2	AAW30256

ALIGNMENTS

RESULT 1  
AAY82576  
ID AAY82576 standard; peptide; 86 AA.  
XX  
AC AAY82576;  
XX  
DT 28-JUL-2000 (first entry)  
XX  
DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:6.  
XX  
KW Copolymer; molecular weight marker; TV-marker; immune disease;  
KW Glutramer acetate; autoimmune disease; antiarthritic; neuroprotective;  
KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;  
KW antidiabetic; thymostatic; haemostatic; antipsoriatic; dermatological;  
KW antiandemic; immunosuppressive; demyelinating disease; osteoarthritis;  
KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;  
KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;  
KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;  
KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;  
KW pemphigus vulgaris; systemic lupus erythematosus.  
XX  
OS Unidentified.  
XX  
FN WO200018794-A1.  
XX  
PD 06-APR-2000.  
XX  
PF 24-SEP-1999; 99WO-US022402.  
XX  
PR 25-SEP-1998; 98US-0101693P.  
XX  
PA (YEDA ) YEDA RES & DEV CO LTD.  
PA (TEVA-) TEVA PHARM USA INC.  
XX  
PI Gad A, Lis D;  
XX  
DR WPI; 2000-317499/27.  
XX  
PT Copolymer 1 related polypeptides used as molecular weight markers for  
PT glutramer acetate and prevention of immune diseases.  
XX  
PS Claim 10; Page 14; 72pp; English.  
XX  
CC AAY82571 to AAY82577 represent specifically claimed copolymer molecular  
CC weight TV-marker polypeptides from the present invention. The present  
CC invention describes polypeptides (I) for determining the molecular weight  
CC of a copolymer (CP), which has an identified molecular weight and an  
CC amino acid composition corresponding to the copolymer. The polypeptides









QY 52 KEYAAAEAKYKAE-----AKKAYKAEAAKAAA---KEAAYEA 86  
: |||:|||||  
Db 183 AK-AAADAKKAEAAEAAKAAAEAKKAEAAEAAKAAAEAKKADA 227

## RESULT 7

ABU44721  
ID ABU44721 standard; protein; 387 AA.

AC ABU44721;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #30248.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Salmonella paratyphi.

OS WO200277183-A2.

PN 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

DR WPI; 2003-029926/02.

DR N-PSDB; ACA48591.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 72645; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 387 AA;

Query Match 40.1%; Score 164; DB 6; Length 387;

Best Local Similarity 57.0%; Pred. NO. 4.9e-07;

Matches 53; Conservative 7; Mismatches 17; Indels 16; Gaps 5;

QY 1 AKYAKKEKAYAKKAEKAAKAEAKYK-AAEAKKAEAKAEAKYK-A---KAAKAEKYE 55  
|||:|||||

Db 141 AKKAEPEA--AKAAADAKKAEAEVKAADAKKAEAEAKAAADAKKAEAEAAK-A 197  
|||:|||||

QY 56 AAEAKYKAE-----AKKAYKAEAAKAAK 80  
|||:|||||

Db 198 AAEAKKAEAEAAKAAADAKKKADAEEAAKAAAE 230  
|||:|||||

## RESULT 8

ABU28559

ID ABU28559 standard; protein; 421 AA.

XX AC ABU28559;

XX 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #14086.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Escherichia coli.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR N-PSDB; ACA32429.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 56483; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)









Job time : 96.0717 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.  
OM protein - protein search, using sw model  
Run on: December 14, 2004, 05:52:22 ; Search time 213.367 Seconds  
(without alignments)  
143.965 Million cell updates/sec

Title: US-10-792-311-6  
Perfect score: 409  
Sequence: 1 AKKYAKKEKAYAKAEKAAK.....KKAYKAEAKAAKAAEAYEA 86

Scoring table: BLASTSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues  
Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	Description
1	409	100.0	86 9 US-09-816-989A-6 Sequence 6, Appli
2	351.5	85.9	77 9 US-09-816-989A-5 Sequence 5, Appli
3	312.5	76.4	109 9 US-09-816-989A-7 Sequence 7, Appli
4	268	65.5	66 9 US-09-816-989A-4 Sequence 4, Appli
5	233	57.0	56 9 US-09-816-989A-3 Sequence 3, Appli
6	165.5	40.5	407 15 US-10-282-122A-75047 Sequence 75047, A
7	164	40.1	387 15 US-10-282-122A-72645 Sequence 72645, A
8	164	40.1	421 15 US-10-282-122A-56483 Sequence 56483, A
9	163	39.9	372 15 US-10-282-122A-56109 Sequence 56109, A
10	161	39.4	428 15 US-10-282-122A-55748 Sequence 55748, A
11	157.5	38.5	323 15 US-10-282-122A-59321 Sequence 59321, A
12	157.5	38.5	376 15 US-10-282-122A-75772 Sequence 75772, A
13	154	37.7	347 14 US-10-127-032-120 Sequence 120, App

14	154	37.7	347	15	US-10-282-122A-66237	Sequence 66237, A
15	148	36.2	389	15	US-10-282-122A-67145	Sequence 67145, A
16	147.5	36.1	388	15	US-10-282-122A-78190	Sequence 78190, A
17	147	35.9	214	14	US-10-229-567-27	Sequence 27, Appl
18	147	35.9	214	15	US-10-282-122A-62547	Sequence 62547, A
19	147	35.9	214	15	US-10-282-122A-64817	Sequence 64817, A
20	142	34.7	372	10	US-09-820-843A-8	Sequence 8, Appli
21	142	34.7	372	16	US-10-467-421-16	Sequence 16, Appli
22	141.5	34.6	336	15	US-10-282-122A-69962	Sequence 69962, A
23	137.5	33.6	448	15	US-10-282-122A-45264	Sequence 45264, A
24	137.5	33.6	507	16	US-10-726-692-50	Sequence 50, Appl
25	137.5	33.6	717	16	US-10-726-692-36	Sequence 36, Appl
26	135	33.0	272	17	US-10-739-930-10710	Sequence 10710, A
27	133.5	32.6	433	15	US-10-390-472-2	Sequence 2, Appli
28	132	32.3	329	15	US-10-282-122A-67699	Sequence 67699, A
29	131	32.0	452	14	US-10-184-832-5	Sequence 5, Appli
30	130.5	31.9	212	15	US-10-282-122A-61735	Sequence 61735, A
31	129.5	31.7	223	13	US-10-051-643-201	Sequence 201, App
32	129.5	31.7	223	14	US-10-205-979-52	Sequence 52, Appli
33	127	31.1	139	15	US-10-282-122A-60257	Sequence 60257, A
34	127	31.1	222	14	US-10-229-567-3	Sequence 3, Appli
35	127	31.1	376	14	US-10-156-761-9889	Sequence 9889, Ap
36	126.5	30.9	45	9	US-09-816-989A-2	Sequence 2, Appli
37	126.5	30.9	568	17	US-10-435-115-197135	Sequence 197135, A
38	126.5	30.9	636	15	US-10-425-114-37076	Sequence 37076, A
39	125	30.6	217	14	US-10-156-761-10221	Sequence 10221, A
40	124.5	30.4	281	16	US-10-767-701-45755	Sequence 45755, A
41	124	30.3	827	16	US-10-437-963-152005	Sequence 152005, A
42	123.5	30.2	926	16	US-10-437-963-193381	Sequence 193381, A
43	123	30.1	130	14	US-10-262-209-2	Sequence 2, Appli
44	123	30.1	130	16	US-10-240-430-5	Sequence 5, Appli
45	123	30.1	356	10	US-09-820-843A-27	Sequence 27, Appli

ALIGNMENTS

RESULT 1

US-09-816-989A-6  
; Sequence 6, Application US/09816989A  
; Patent No. US20020115103A1  
; GENERAL INFORMATION:  
; APPLICANT: Gad, Alexander  
; APPLICANT: Lis, Doris  
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK  
; FILE REFERENCE: 2609/60807-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/816,989A  
; PRIOR FILING DATE: 2001-03-23  
; PRIOR FILING DATE: 1998-09-25  
; PRIOR APPLICATION NUMBER: PCT/US99/22402  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 6  
; LENGTH: 86  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide  
US-09-816-989A-6

Query Match				100.0%;	Score 409;	DB 9;	Length 86;
Best Local Similarity				100.0%;	Pred. No. 4.3e-28;		
Matches				86;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	AKKYAKKEKAYAKAEKAAKAEKAYKAAEAKKKAKAEKKYAKAAKAEKKEYYAAAEAK	60				
Db	1	AKKYAKKEKAYAKAEKAAKAEKAYKAAEAKKKAKAEKKYAKAAKAEKKEYYAAAEAK	60				
Qy	61	YKAEAAKKYKAEAAKAAKAAEAYEA	86				









```
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55748
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55748

Query Match      39.4%; Score 161; DB 15; Length 428;
Best Local Similarity 47.5%; Pred. No. 4e-06;
Matches 48; Conservative 16; Mismatches 21; Indels 16; Gaps 3;

QY 1 AKYAKKEKAYAKAEKAKK-----AEAKYKAAAEKKAKEAKYKAAKAEKK 52
   |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 127 ABEAKKQAEQEQKQAEAEAKAAADAKAQADQAKLAFAEAKKAAADAKQKAE-AEAKK 185
   |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 53 EYAAAEAKYKAEAKKA-----YKAEAAKAAAEAEAYEA 86
   |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 186 AAADAKQKAEAEAKKAAADAKQKAEAEAKKAAQAEKKA 226
   |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

RESULT 11
US-10-282-122A-59321
; Sequence 59321, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09

; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59321
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59321

Query Match      38.5%; Score 157.5; DB 15; Length 323;
Best Local Similarity 55.7%; Pred. No. 5.9e-06;
Matches 44; Conservative 11; Mismatches 21; Indels 3; Gaps 2;

QY 8 EKAVAKAEKAAKAEAKAYK-AEAKKKAAAEAKKYAKAAKAEKKEYYAAAEAKYAEAA 66
   |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 55 EQAAKAAADAKQAEAAAKAAAEAKKQAEAEAAK--AAAEAKKAEAAAKKQAEAE 112
   |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 67 KKAYKAEAKAAKEAEAYE 85
   |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 113 KKAQAEAAKQAAAEKAAAE 131

RESULT 12
US-10-282-122A-75772
; Sequence 75772, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
```



```
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatenIn version 3.1
; SEQ ID NO 67145
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-282-122A-67145

Query Match      36.2%; Score 148; DB 15; Length 389;
Best Local Similarity 47.2%; Pred.No. 4.7e-05;
Matches 50; Conservative 11; Mismatches 23; Indels 22; Gaps 5;

QY      2 KKYAKKEKAYAKAEKAAK-KAEAKAYKAA-----EAKKKAKAEAKKYAKAKAE-- 50
Db      142 EKQKQAEAKAKQLAEAAKKAEEAKRUAALAKQAEAEKAKAAEEAKKAEKAEKAE 201

QY      51 -----KKEYAAAEAKYKAEAAK-----KAYKAEAAKAAKAEAAEYEA 86
Db      202 AKAKVEKAKAEAEAKVKAEKAEAEAKA-KAEKAKAEAKAEKAEKAE 246
```

Search completed: December 14, 2004, 06:59:42  
Job time : 214.367 secs

***This Page Blank (uspio)***

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	409	100.0	86	4	US-09-405-743A-6	Sequence 6, Appli
2	351.5	85.9	77	4	US-09-405-743A-5	Sequence 5, Appli
3	312.5	76.4	109	4	US-09-405-743A-7	Sequence 7, Appli
4	268	65.5	66	4	US-09-405-743A-4	Sequence 4, Appli
5	233	57.0	56	4	US-09-405-743A-3	Sequence 3, Appli
6	157.5	38.5	469	4	US-09-489-039A-13565	Sequence 13565, A
7	156	38.1	100	2	US-08-460-890A-64	Sequence 64, Appl
8	136	38.1	100	3	US-08-167-641C-64	Sequence 64, Appl
9	156	38.1	100	3	US-08-460-971A-64	Sequence 64, Appl
10	156	38.1	100	3	US-08-462-040-64	Sequence 64, Appl
11	154	37.7	407	4	US-09-252-991A-29581	Sequence 29581, A
12	147	35.9	214	3	US-09-041-889-27	Sequence 27, Appl
13	147	35.9	214	4	US-09-417-264-27	Sequence 27, Appl
14	141.5	34.6	361	4	US-09-543-681A-5390	Sequence 5390, Ap
15	140.5	34.4	214	4	US-09-328-352-5169	Sequence 5169, Ap
16	140	34.2	468	4	US-09-328-352-6321	Sequence 6321, Ap
17	137.5	33.6	472	2	US-08-216-894-10	Sequence 10, Appl
18	137.5	33.6	472	3	US-09-115-746-10	Sequence 10, Appl
19	137.5	33.6	564	2	US-08-216-894-2	Sequence 2, Appli
20	137.5	33.6	564	3	US-09-115-746-2	Sequence 2, Appli
21	137.5	33.6	643	3	US-08-216-894-8	Sequence 8, Appli
22	137.5	33.6	643	3	US-09-115-746-8	Sequence 8, Appli
23	133.5	32.6	433	1	US-08-346-849-2	Sequence 2, Appli
24	133.5	32.6	433	2	US-08-293-284A-2	Sequence 2, Appli
25	133.5	32.6	433	4	US-08-898-300-2	Sequence 2, Appli
26	129.5	31.7	223	3	US-09-095-855-201	Sequence 201, App
27	129.5	31.7	223	4	US-09-205-426-201	Sequence 201, App





Db	205	EQAAKAAADAKKQAEAAAKAAAEAKKQAEAAK--AAAEQKKAEEAAAKKAAQEEAE	262
Qy	67	KKRYKAEAAKAAAEAAE	85
		:   :     :	
Db	263	KKAQEEAAKQAAAEKAAAE	281

RESULT 7  
US-08-460-890A-64  
; Sequence 64, Application US/08460890A  
; Patent No. 5994109

/ GENERAL INFORMATION:  
 / APPLICANT: Woo, Savio L.C.  
 / APPLICANT: Smith, Louis C.  
 / APPLICANT: Cristiano, Richard J.  
 / APPLICANT: Gotchalk, Stephen  
 / TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND  
 / TITLE OF INVENTION: METHODS OF USE  
 / NUMBER OF SEQUENCES: 65  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: Lyon & Lyon  
 / STREET: 633 West Fifth Street  
 / STREET: Suite 4700  
 / CITY: Los Angeles  
 / STATE: California  
 / COUNTRY: U.S.A.  
 / ZIP: 90071-2066

```

;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;
; MEDIUM TYPE: storage
;
; COMPUTER: IBM Compatible
;
; OPERATING SYSTEM: IBM P.C. DOS 5.0
;
; SOFTWARE: FastSeq for Windows 2.0
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/460,890A
;
; FILING DATE: June 5, 1995
;

```

```

: CLASSIFICATION: 435
:
: PRIOR APPLICATION DATA:
:
: APPLICATION NUMBER: 08/167,641
: FILING DATE: December 14, 1993
: APPLICATION NUMBER: 07/855,389
: FILING DATE: March 20, 1992
: APPLICATION NUMBER: PCT/US93/02725
: FILING DATE: March 19, 1993
:
: ATTORNEY/AGENT INFORMATION:
:
: NAME: Warburg, Richard J.
: REGISTRATION NUMBER: 32,327
: REFERENCE/DOCKET NUMBER: 212/066
:
: TELECOMMUNICATION INFORMATION:
:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELE: 67-3510
:

```

```

;
; INFORMATION FOR SEQ ID NO: 64:
;
; SEQUENCE CHARACTERISTICS:
;
;   LENGTH: 100 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;
; MOLECULE TYPE: peptide
;
; FEATURE:
;

```

; OTHER INFORMATION: "Lys Ala" in positions 3 to 100 may be  
 ; OTHER INFORMATION: present or absent.  
 US-08-460-890A-64

Query Match 38.1%; Score 156; DB 2; Length 100;  
Best Local Similarity 55.3%; Pred. No. 7.7e-08;  
Matches 47; Conservative 6; Mismatches 30; Indels

[illegible]

QY 59 AKYKAEAAKKAYKAEAAKAAAKEAA 833

[illegible]

RESULT 8

```

US-08-167-641C-64
; Sequence 64, Application US/08167641C
; Patent No. 6033884
; GENERAL INFORMATION:
; APPLICANT: Woo, Savio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gottchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066

```

```

/
/ COMPUTER READABLE FORM:
/
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
/
/ MEDIUM TYPE: storage
/
/ COMPUTER: IBM Compatible
/
/ OPERATING SYSTEM: IBM P.C. DOS 5.0
/
/ SOFTWARE: FASTSEQ for Windows 2.0
/
/ CURRENT APPLICATION DATA:
/
/ APPLICATION NUMBER: US/08/167,641C
/
/ FILING DATE: December 14, 1993
/
/ CLASSIFICATION: 435
/
/

```

; PRIOR APPLICATION DATA: 07/855,189  
 ; APPLICATION NUMBER: 07/855,189  
 ; FILING DATE: March 20, 1992  
 ; APPLICATION NUMBER: PCT/US93/01  
 ; FILING DATE: March 19, 1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Warburg, Richard J.  
 ; REGISTRATION NUMBER: 32,327  
 ; REFERENCE/DOCKET NUMBER: 205/01

```
, , TELECOMMUNICATION INFORMATION:
, , TELEPHONE: (213) 489-1600
, , TELEFAX: (213) 955-0440
, , TELEX: 67-3510
, , INFORMATION FOR SEQ ID NO: 64:
, , SEQUENCE CHARACTERISTICS:
, , LENGTH: 100 amino acids
, , TYPE: amino acid
, , STRANDEDNESS: single
, , TOPOLOGY: linear
, , MOLECULE TYPE: peptide
, , FEATURE:
```

OTHER INFORMATION: "Lys Ala" in positions 3 to 100 may be present or absent.

US-08-167-641C-64

Query Match 38.1%; Score 156; DB 3; Length 100;  
Best Local Similarity 55.3%; Pred. No. 7.7e-08;  
Matches 47; Conservative 6; Mismatches 30; Indels

[illegible]

QY 59 AKYKAEAAKKAYKAEAAKAAKEAA 83

Db 62 AKAKAKAKAKAKAKAKAKAKAKA 86

## RESULT 9



; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29581
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29581

Query Match 37.7%; Score 154; DB 4; Length 407;
Best Local Similarity 47.6%; Pred. No. 5e-07;
Matches 49; Conservative 11; Mismatches 21; Indels 22; Gaps 4;

QY 1 AKKYAKKERAKYAKAEAKKAEAKYKAA-----EAKKAKAEAKKYA---- 44
Db 180 ARKAEQAQAAEAKKAEAKAEAKAEQKQADIAKRAEAEAKKAAEDAKKAAEDA 239
QY 45 --KAAYAEKEVAAAEAKYKA--EAKKAYKAEAAKAAKAA 83
Db 240 KKAAAEAKKAAAEAAKKAAYEAAK--KAAAEAAAKKAA 280

RESULT 12
US-09-041-889-27
; Sequence 27, Application US/09041889
; Patent No. 6033864
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; APPLICANT: Cohavy, Offer
; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of
; TITLE OF INVENTION: Ulcerative Colitis, and Clinical Subtypes Thereof, Using
; TITLE OF INVENTION: Microbial UC PANCA antigens
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,889
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/837,058
; FILING DATE: 11-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-PM 3006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-041-889-27

Query Match 35.9%; Score 147; DB 3; Length 214;
Best Local Similarity 50.0%; Pred. No. 1.1e-06;
Matches 47; Conservative 3; Mismatches 34; Indels 10; Gaps 3;

QY 1 AKKYAKKERAKYAKAEAKKAEAK-----AYKAAEAKKKAKAEAKKYAKAAKAEKKEY 54
Db 111 AKKVAK--KAPAKKATKAAKKAATKAPAKKAATKAPAKKAATKAPAKKAVKATKSPAKKV 168
QY 55 AAAEAK--YKAEAAKKAYKAEAAKAAKAAEAAVEA 86
Db 169 TKAVKKTAVKASVRKAATKAPAKKAAAKRPATKA 202

RESULT 13
US-09-417-264-27
; Sequence 27, Application US/09417264
; Patent No. 6537768
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; APPLICANT: Cohavy, Offer
; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of
; TITLE OF INVENTION: Ulcerative Colitis, and Clinical Subtypes Thereof, Using
; TITLE OF INVENTION: Microbial UC PANCA antigens
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/417,264
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/041,889
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-PM 3006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-417-264-27

Query Match 35.9%; Score 147; DB 4; Length 214;
Best Local Similarity 50.0%; Pred. No. 1.1e-06;
Matches 47; Conservative 3; Mismatches 34; Indels 10; Gaps 3;

QY 1 AKKYAKKERAKYAKAEAKKAEAK-----AYKAAEAKKKAKAEAKKYAKAAKAEKKEY 54
Db 111 AKKVAK--KAPAKKATKAAKKAATKAPAKKAATKAPAKKAATKAPAKKAVKATKSPAKKV 168
QY 55 AAAEAK--YKAEAAKKAYKAEAAKAAKAAEAAVEA 86
Db 169 TKAVKKTAVKASVRKAATKAPAKKAAAKRPATKA 202

RESULT 14
US-09-543-681A-5390

```
; Sequence 5390, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5390
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5390

Query Match      34.6%; Score 141.5; DB 4; Length 361;
Best Local Similarity 52.1%; Pred. No. 6.3e-06;
Matches 50; Conservative 12; Mismatches 19; Indels 15; Gaps 6;

QY      5 AKKEKAYAKAEAKKAEAK-----AYKAEAKKAKAEAKKYAKAA---KAEK-KEYAA 56
Db      138 AAKAKSEQQAEEAAQAQAEADRILKEQADAKAKAEAEAKQAELAAKQKAEAEKAKAE 197

QY      57 AEAKYKAEAAKAYKAEAE---AKAAAK---EAAVEA 86
Db      198 AEAKAEADAKA-KAEADAKAKAEAEKAKAAAEAA 232

RESULT 15
US-09-328-352-5169
; Sequence 5169, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5169
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5169

Query Match      34.4%; Score 140.5; DB 4; Length 214;
Best Local Similarity 51.0%; Pred. No. 4.5e-06;
Matches 50; Conservative 10; Mismatches 23; Indels 15; Gaps 6;

QY      1 AKKYAKKEK--AVAKKAEKAAKAEAKYKAAEAKKYKAAEAKKYKAAK--AEKKEYA- 55
Db      29 AKRLAEADKKAAEAKRQAEADKKA-AEAKRQAEADKKA-AEAKRQAEADKKAAEAKRQAE 86

QY      56 ----AEAEKYKAEAKKA-----YKAEAAKAAAEAAVE 85
Db      87 ADKKAAEAKRQAEADKKAAEAKRKAEEAKKAEAEKARE 124

Search completed: December 14, 2004, 05:50:18
Job time : 24.2737 secs
```

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 25.0654 Seconds  
(without alignments)  
418.411 Million cell updates/sec

Title: US-10-792-311-7  
Perfect score: 519  
Sequence: 1 AKKYAKKAEKAYAKKAAK.....AKAYKAEAAKAAKAEAYEA 109

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 segs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79: \*  
1: Piri: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	189.5	36.5	421	JV0057	tola protein - Esc
2	181.5	35.0	394	F90725	membrane spanning
3	181.5	35.0	394	G85576	membrane spanning
4	181	34.9	210	A25550	histone H1 - sea u
5	179	34.5	347	E83525	Tola protein PA097
6	178	34.3	376	AG0592	Tola protein [impo
7	178	34.3	388	AC0138	Tola colicin impor
8	175.5	33.8	344	S34153	mst101-1 protein -
9	172.5	33.2	211	A28100	histone H1-beta, e
10	171.5	33.0	1390	S51364	sperm tail-specifi
11	170.5	32.9	248	HSUR1P	histone H1, gonada
12	170.5	32.9	1701	T09127	probable erythrocy
13	165	31.8	220	A28456	histone H1.10 - ch
14	164.5	31.7	231	S59589	histone H1 - Chlam
15	164	31.6	265	S19113	cgr-4 protein - C
16	164	31.6	311	T17698	hypothetical prote
17	163.5	31.5	384	B43592	outer membrane pro
18	163	31.4	182	S61926	histone H1 homolog
19	163	31.4	225	B28456	histone H1.11L - c
20	162.5	31.3	1128	T30296	R27-2 protein - Tr
21	162	31.2	356	A82152	tola protein VC193
22	161.5	31.1	328	A44993	cytosolic repetiti
23	161	31.0	219	E60110	repetitive protein
24	160.5	30.9	206	HSR1R	histone H1 - rainb
25	160.5	30.9	226	S51660	histone H1-5 [vali
26	159.5	30.7	219	C28456	histone H1.11R - c
27	158	30.4	219	HSR1B	histone H1-4 [vali
28	157	30.3	771	A33430	h-caldesmon - chic
29	156.5	30.2	218	A23055	histone H1.01 - ch

30	154.5	29.8	217	2	S29309	hypothetical prote
31	154.5	29.8	241	2	JN0748	histone H1-II - Vo
32	154.5	29.8	309	2	G83013	polyhydroxyalkanoa
33	153.5	29.6	369	2	H71321	conserved hypothet
34	153.5	29.6	372	2	G64064	outer membrane int
35	152.5	29.4	224	2	D28456	histone H1.03 - ch
36	151	29.1	1403	2	T11583	probable translati
37	150	28.9	212	2	A28470	histone H1 - mouse
38	150	28.9	217	2	A26721	histone H1 - gamma,
39	150	28.9	288	2	T06257	histone H1 (clone
40	148	28.5	218	2	S01262	histone H1 - musco
41	147.5	28.4	219	2	I49742	histone H1 - mouse
42	147	28.3	208	2	T23778	histone H1.1 - Cae
43	147	28.3	221	2	S49482	hypothetical prote
44	147	28.3	243	2	AE1689	hypothetical prote
45	147	28.3	340	2	A35630	regulatory protein

ALIGNMENTS

RESULT 1

JV0057  
tola protein - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004  
C:Accession: JV0057; B64810  
R:Levensgood, S.K.; Webster, R.E.  
J: Bacteriol. 171, 6600-6609, 1989  
A:Title: Nucleotide sequences of the tola and tolB genes and localization of their produc  
A:Reference number: JV0057; MUID:90078104; PMID:2687247  
A:Accession: JV0057  
A:Molecule type: DNA  
A:Residues: 1-421 <LEV>  
A:Cross-references: UNIPROT:P19934; GB:M28232; NID:G148018; PIDN:AAA24683.1; PID:G148019  
A:Experimental source: strain JM105  
A>Note: The authors translated the initiation codon GTG for residue 1 as Val  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Coi  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: B64810  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-421 <BLAT>  
A:Cross-references: GB:AE000177; GB:U00096; NID:G1786955; PIDN:AAC73833.1; PID:G1786960;  
A:Experimental source: strain K-12, substrain MG1655  
C:Comment: tola and tolB proteins are necessary for colicins E2, E3, A, and K to reach th  
C:Genetics:  
A:Gene: tola  
A:Map position: 17 min  
A:Start codon: GTG  
C:Keywords: nucleotide binding; P-loop; transmembrane protein  
F:14-34/Domain: transmembrane #status predicted <MSS>  
F:78-301/Domain: helical #status predicted <HSR>  
F:355-362/Region: nucleotide-binding motif A (P-loop)

Query Match	36.5%;	Score	189.5;	DB 2;	Length	421;
Best Local Similarity	53.0%;	Pred. No.	1.7e-05;			
Matches	61;	Conservative	11;	Mismatches	32;	Indels 11; Gaps 5;
Qy	1	AKKYAKKAE-----KAYAKKAAKAEKAYAKKEKAYKAAEAKKKAKAEAKKYAK-EAAK	55			
Db	120	AEAAKQAELEKQKQAEAAAKAAADAKAAEADAKA--AEAAKKAADAKKAAEAAK	177			
Qy	56	AKKEAY-KAEAKKYAKAAKAEKKEYYAAAEAKK---AEAAKAYKAEAAKAAAEKAA	106			
Db	178	AAAEAKQKAAEAAAALKKAAEAAAEAAAEAAKKAATEAAEKAAEAKKAAAEKA	232			

RESULT 2

F90725

```
membrane spanning protein Tola [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F90725
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90725
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-394 <HAY>
A;Cross-references: UNIPROT:Q8X965; GB:BA000007; PIDN:BAB34197.1; PID:g13360233; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 050952
C;Genetics:
A;Gene: EC60774

Query Match 35.0%; Score 181.5; DB 2; Length 394;
Best Local Similarity 52.2%; Pred. No. 5.3e-05;
Matches 60; Conservative 12; Mismatches 32; Indels 11; Gaps 5;

QY 1 AKKYAKAE-----KAYAKKAAKEKKYAKYKAEKKAAYKAAEKKKAKAEAKKYAK-BAAK 55
Db 120 AEEAAKQAEIKQKQAEBAAKAAADAKAKAEADDKA--ABEAAKKAADAKKKAEBAAK 177
QY 56 AKKEAY-KAEAKYKAAKAEKKEYYAAAEAKK---AEAAKAYKAEAAKAAKEAA 106
Db 178 AAABEQKAAEAAALKKKAAEAAEAAAEARKKAAAEKAAADKKAKEKAAAEKAA 232

RESULT 3
G85576
membrane spanning protein Tola [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: G85576
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: G85576
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-394 <STO>
A;Cross-references: UNIPROT:Q8X965; GB:AE005174; NID:gl2513672; PIDN:AGS5075.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: tola

Query Match 35.0%; Score 181.5; DB 2; Length 394;
Best Local Similarity 52.2%; Pred. No. 5.3e-05;
Matches 60; Conservative 12; Mismatches 32; Indels 11; Gaps 5;

QY 1 AKKYAKAE-----KAYAKKAAKEKKYAKYKAEKKAAYKAAEKKKAKAEAKKYAK-BAAK 55
Db 120 AEEAAKQAEIKQKQAEBAAKAAADAKAKAEADDKA--ABEAAKKAADAKKKAEBAAK 177
QY 56 AKKEAY-KAEAKYKAAKAEKKEYYAAAEAKK---AEAAKAYKAEAAKAAKEAA 106
Db 178 AAABEQKAAEAAALKKKAAEAAEAAAEARKKAAAEKAAADKKAKEKAAAEKAA 232

RESULT 4
A25550
histone H1 - sea urchin (Lytechinus pictus)
C;Species: Lytechinus pictus (painted urchin)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
C;Accession: A25550
R;Knowles, J.A.; Childs, G.J.
Nucleic Acids Res. 14, 8121-8133, 1986
A;Title: Comparison of the late H1 histone genes of the sea urchins Lytechinus pictus an
```

```
A;Reference number: A25550; MUID:87040778; PMID:3022245
A;Accession: A25550
A;Molecule type: DNA
A;Residues: 1-210 <KNO>
A;Cross-references: UNIPROT:P06144; GB:X04488; NID:g9616; PIDN:CAA28177.1; PID:g9617
C;Superfamily: histone H1
C;Keywords: chromosomal protein; DNA binding; nucleosome; nucleus

Query Match 34.9%; Score 181; DB 2; Length 210;
Best Local Similarity 50.0%; Pred. No. 3.6e-05;
Matches 58; Conservative 8; Mismatches 36; Indels 14; Gaps 4;

QY 2 KKYAKAEKAYAKKAAKAEKKYAKYKAEKKAAYKAAE----AKKKAKAEAKKYAKEAAKAK 57
Db 96 KTEAQKA-RAAAKKAALAAKKEQKEKKAATKARKEKLAAKAAKAAKKAKKVKKPAAKAK 154
QY 58 KEAYKAEAKYKAAKAEKKEYYAAAEAKKA-----EAAKAYKAEAAKAAAEKAA 106
Db 155 KPAKKAACKPAAK--KAACKPAACKPAKKAACKPAKKAACKPAKKAACKPAKKAACKAA 208

RESULT 5
E83525
Tola protein PA0971 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: E83525
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathos
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83525
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-347 <STO>
A;Cross-references: UNIPROT:P50600; GB:AE004530; GB:AE004091; NID:g9946865; PIDN:AAG0436
A;Experimental source: strain PA01
C;Genetics:
A;Gene: tola; PA0971

Query Match 34.5%; Score 179; DB 2; Length 347;
Best Local Similarity 49.1%; Pred. No. 6.9e-05;
Matches 54; Conservative 17; Mismatches 29; Indels 10; Gaps 4;

QY 6 KKAEEKAYAKKAAAEK-----KAYKKEAKYKAAEAKKKAKEAKKYKAEAKKAKE 59
Db 99 OKLEQQVAAAKAAEQKKADEARKAEQAQKAAEAKKAAEAKKAAEAKAAEQKKQADIKKR 158
QY 60 AYKAEAKYKAAKAEKKEYYAAAEAKKAEEAAKAYKAEAAKAAKAAAEAA 109
Db 159 A-EDEAKK--KAEDAKKK-AAEDAKKKAEEAKKKAEEAAKAAAEAAKKAEEAA 204

RESULT 6
AG0592
tola protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain CT18)
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AG0592
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
; S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serova
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AG0592
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-376 <PAR>
```



```
QY 1 AKKYAKAEKAY-----AKKAKAAEKKAYAKAEKAYKAAE-----AKKKAKAEAKKY 49
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1067 AKKLKEAAEKQCEERAKKEAEAKKQCEERAKKLKEAAEKQCEERAKKEAEAKK 1126

QY 50 AKEAAKAEKAYKAEAKKAYAKAEKAYKAEKAYKAEKAYKAEKAYKAEKAYKAEKAY 106
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1127 CEEAAKREKEA--AEKKCAEAAKKE--ATEKQKCAEAAKKEAEKAEKKCAEAA 1179

RESULT 11
HSUR1P
histone H1, gonadal - sea urchin (Parechinus angulosus)
C:Species: Parechinus angulosus (langulate urchin)
C>Date: 31-Mar-1980 #sequence_revision 11-Mar-1980 #text_change 09-Jul-2004
C:Accession: A91090; A91091; A02586
R:Strickland, W.N.; Strickland, M.; de Groot, P.C.; von Holt, C.; Wittmann-Liebold, B.
Eur. J. Biochem. 104, 559-566, 1980
A:title: The primary structure of histone H1 from sperm of the sea urchin Parechinus angulosus
A:Reference number: A91090; MUID:80156831; PMID:6767609
A:Contents: sequence of residues 1-84
A:Accession: A91090
A:Molecule type: protein
A:Residues: 1-248 <STR>
A:Cross-references: UNIPROT:P02256
R:Strickland, W.N.; Strickland, M.; Brandt, W.F.; von Holt, C.; Lehmann, A.; Wittmann-Liebold, B.
Eur. J. Biochem. 104, 567-578, 1980
A:title: The primary structure of histone H1 from sperm of the sea urchin Parechinus angulosus
A:Reference number: A91091; MUID:80156832; PMID:7363905
A:Accession: A91091
A:Molecule type: protein
A:Residues: 80-248 <ST2>
A:Note: 144-Arg was also found
C:Superfamily: histone H1
C:Keywords: DNA binding; nucleosome; sperm

Query Match 32.9%; Score 170.5; DB 1; Length 248;
Best Local Similarity 46.0%; Pred. No. 0.00018;
Matches 52; Conservative 9; Mismatches 41; Indels 11; Gaps 2;

QY 2 KKYAKAEKAYKAAKAEKKAYAKAEKAYKAAKAEKAYKAAKAEKAYKAAKAEKAYKAA 61
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 120 KKAKTSAAKAKKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAK 179

QY 62 KA--EAKKYAK-----AAKAEKAYAAAEKAEKAEKAYKAAKAEKAYKAAKAAK 103
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 180 KAAKAKKPAKSPKAKKPAKSPKAKKSPKAKKSPKAKKSPKAKKSPKAKKSPKAKK 232

RESULT 12
T09127
probable erythrocyte-binding protein MABBL - Plasmodium yoelii
C:Species: Plasmodium yoelii
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T09127
R:Kapke, S.H.I.; Noe, A.R.; Fraser, T.S.; Blair, P.L.; Adams, J.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 1230-1235, 1998
A:title: A family of chimeric erythrocyte binding proteins of malaria parasites.
A:Reference number: Z16577; MUID:98115903; PMID:9448314
A:Accession: T09127
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1701 <KAP>
A:Cross-references: UNIPROT:O61164; EMBL:AF031886; NID:G2947227; PID:G2947228
A:Experimental source: subspecies yoelii; strain YM
C:Genetics:
A:Gene: mabbl
A:Introns: 62/1; 1648/1; 1674/2; 1697/1
C:Keywords: alternative splicing; cell binding; erythrocyte invasion

Query Match 32.9%; Score 170.5; DB 2; Length 1701;
Best Local Similarity 53.5%; Pred. No. 0.00076;
```

```
Matches 54; Conservative 11; Mismatches 31; Indels 5; Gaps 5;
QY 5 AKKAEKAYKAAKAEKKAYAKAEKAYKAAKAEKKAYKAAKAEKKAYKAAKAEKKAYKAA 63
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1205 AKKAEKE--RKGAENV-KAEAEKKAEKAEKAEKAEKK-KAEAEKAEKAEKKAEKAEK 1261

QY 64 EAKYKAAKAEKAYKAAKAEKKAYKAAKAEKKAYKAAKAEKKAYKAAKAEKKAYKAA 104
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1262 ERKKAEKAEKAEKKAEKAEKKAEKAEKK-KAEAEKAEKAEKEE 1301

RESULT 13
A28456
histone H1.10 - chicken
C:Species: Gallus gallus (chicken)
C>Date: 28-Aug-1989 #sequence_revision 28-Aug-1989 #text_change 04-Mar-2000
C:Accession: A28456
R:Colles, L.S.; Robins, A.J.; Madley, L.K.; Wells, J.R.E.
J. Biol. Chem. 262, 9656-9663, 1987
A:title: Characterization of the chicken histone H1 gene complement. Generation of a complementary DNA library
A:Reference number: A92655; MUID:87250632; PMID:3597432
A:Accession: A28456
A:Molecule type: DNA
A:Residues: 1-220 <COL>
A:Cross-references: GB:M17018; NID:G211834; PIDN:AAA48788.1; PID:G211835
C:Superfamily: histone H1
C:Keywords: acetylated amino end; chromosomal protein; DNA binding; nucleosome; nucleus
P;2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted

Query Match 31.8%; Score 165; DB 2; Length 220;
Best Local Similarity 46.4%; Pred. No. 0.00038;
Matches 52; Conservative 11; Mismatches 41; Indels 8; Gaps 3;

QY 1 AKKYAKAEKAYKAAKAEKKAYKAAKAEKKAYKAAKAEKKAYKAAKAEKKAYKAAKAE 53
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 109 SKKPGVEKKEPKRTPPAKPKPAKPKPAKPAKPAKPAKPAKPAKPAKPAKPAKPA 168

QY 54 AKAEKAEKAYKAAKAEKKAYKAAKAEKKAYKAAKAEKKAYKAAKAEKKAYKAAKAE 104
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 169 AKSPKATYKAAKPKKAAKTAAPKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAK 220

RESULT 14
S59589
histone H1 - Chlamydomonas reinhardtii
C:Species: Chlamydomonas reinhardtii
C>Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C:Accession: S59589; S62122
R:Fabry, S.; Mueller, K.; Lindauer, A.; Park, P.B.; Corneliussen, T.; Schmitt, R.
Curr. Genet. 28, 333-345, 1995
A:title: The organization structure and regulatory elements of Chlamydomonas histone gene
A:Reference number: S59581; MUID:96120862; PMID:8590479
A:Accession: S59589
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-231 <PAB>
A:Cross-references: UNIPROT:Q39576; EMBL:U16726
A:Note: the authors did not translate the codon for residue 1
R:Fabry, S.; Ford, C.; Lindauer, A.; Mueller, K.; Corneliussen, T.; Schmitt, R.
submitted to the EMBL Data Library, October 1994
A:Description: The organization, structure and controlling elements of Chlamydomonas histone gene
A:Reference number: S62122
A:Accession: S62122
A:Molecule type: DNA
A:Residues: 1-173; P', 174-231 <FAW>
A:Cross-references: EMBL:U16726; NID:G571479; PIDN:AAA98452.1; PID:G571480
C:Genetics:
A:Introns: 62/3; 101/3
C:Superfamily: histone H1
C:Keywords: chromosomal protein; DNA binding; nucleosome

Query Match 31.7%; Score 164.5; DB 2; Length 231;
Best Local Similarity 45.9%; Pred. No. 0.00042;
```





***This Page Blank (uspto)***

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 122.338 Seconds  
(without alignments)  
512.646 Million cell updates/sec

Title: US-10-792-311-7  
Perfect score: 519  
Sequence: 1 AKYAKAEKAYAKAKAAK.....AKAYKAEAAKAAKAAAYEA 109

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_02: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	202	38.9	395	2	Q937K4	Q937k4 erwinia chr
2	193.5	37.3	1701	2	Q7RC08	Q7rc08 plasmodium
3	190.5	36.7	372	2	Q9WWX1	Q9wwx1 pseudomonas
4	190.5	36.7	372	2	Q8BN16	Q8bn16 pseudomonas
5	190.5	36.7	1817	2	Q7K5Q8	Q7k5q8 plasmodium
6	190.5	36.7	1817	2	AAQ73467	AAq73467 plasmodium
7	190.5	36.7	1830	2	Q7K5Q7	Q7k5q7 plasmodium
8	190.5	36.7	1830	2	AAQ73466	AAq73466 plasmodium
9	190.5	36.7	1866	2	Q8T5C8	Q8t5c8 plasmodium
10	189.5	36.5	421	1	TOLA ECOLI	P19934 escherichia
11	188.5	36.3	421	2	Q8EJ71	Q8fjt1 escherichia
12	186.5	35.9	401	2	Q74W64	Q74w64 yersinia pe
13	186.5	35.9	401	2	AA561283	AA561283 yersinia
14	184.5	35.5	407	2	Q8ZQT6	Q8zqt6 salmonella
15	181.5	35.0	394	2	Q8XG18	Q8agi8 escherichia
16	181.5	35.0	394	2	Q8X965	Q8x965 escherichia
17	181	34.9	210	1	H1 LYTP1	P06144 lytechinus
18	179	34.5	347	1	TOLA_PSEAE	P50600 pseudomonas
19	179	34.5	395	2	Q6D7F3	Q6d7f3 erwinia car
20	178	34.3	376	2	Q8Z8C1	Q8z8c1 salmonella
21	178	34.3	388	2	Q8ZGZ2	Q8zgzt yersinia pe
22	178	34.3	393	2	Q8CZ28	Q8czs8 yersinia pe
23	177.5	34.2	420	2	Q73D27	Q73d27 bacillus ce
24	177.5	34.2	420	2	AA539818	AA539818 bacillus
25	177	34.1	713	2	Q6CDX0	Q6cdx0 yarrowia li
26	176.5	34.0	200	2	Q8XVW7	Q8xvn7 ralestonia s
27	176.5	34.0	413	2	Q8C2Q4	Q7c2q4 shigella fl
28	176.5	34.0	413	2	Q83SA1	Q83eal shigella fl
29	176	33.9	177	2	Q6SG84	Q6sg84 uncultured
30	176	33.9	177	2	AA37978	AA37978 unculture
31	176	33.9	1268	2	Q6CGN4	Q6cgn4 yarrowia li

32	175.5	33.8	344	1	MST1_DROHY	Q08695 drosophila
33	173.5	33.4	389	2	Q9CM70	Q9cm70 pasteurella
34	173.5	33.4	524	2	Q6HHE7	Q6hhe7 bacillus th
35	171.5	33.0	1391	1	MST2_DROHY	Q08696 drosophila
36	171	32.9	436	2	Q6F986	Q6f986 acinetobact
37	170.5	32.9	248	1	H1 PARAN	P02256 parechinus
38	170.5	32.9	1652	2	Q7KP21	Q7kpz1 plasmodium
39	170.5	32.9	1652	2	AAQ73456	AAq73456 plasmodium
40	170.5	32.9	1680	2	Q7KPY9	Q7kpy9 plasmodium
41	170.5	32.9	1680	2	AAQ73455	AAq73455 plasmodium
42	170.5	32.9	1680	2	AAQ73457	AAq73457 plasmodium
43	170.5	32.9	1701	2	O61164	O61164 plasmodium
44	170	32.8	232	2	Q39576	Q39576 chlamydomon
45	169.5	32.7	211	1	H1B_STRPU	P15869 strongyloce

ALIGNMENTS

RESULT 1

Q937K4

ID Q937K4 PRELIMINARY; PRT; 395 AA.

AC Q937K4; (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DE Tola protein.

DE Tola protein.

GN Name=tola;

OS Erwinia chrysanthemi.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Pectobacterium.

OX NCBI\_TaxID=556;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=3937;

RA Ray M.C., Vianney A., Cotte-pattat N., Lazzaroni J.;

RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ297885; CAC82708.1; -.

DR HSSP; P19934; ITOL.

DR InterPro; IPR010528; Tola.

DR Pfam; PF06519; Tola; 1.

SQ SEQUENCE 395 AA; 41601 MW; 3C0C1DC12E181013 CRC64;

Query Match 38.9%; Score 202; DB 2; Length 395;

Best Local Similarity 47.3%; Pred. No. 6.5e-06;

Matches 62; Conservative 15; Mismatches 32; Indels 22; Gaps 4;

QY 1 AKYAKAEKAYAK-----KAKAEKAYAKAEKAYAKAEKAKKAK-----43

DB 138 AKEQOKQAEVAAAKAKAEAEQQAQAAADAKKQAEVEVKQAADAKKQAEAEAKTKYKAAA 197

QY 44 AEAKKYAKAEAKKKEA---YKAEAKYAKAAKAEKKEVYAAAEAKKAEAAKAYK--AEAA 98

DB 198 AEAKKAEAEKAKAAADAKKQAEAEAKKAEKAEKAAADAAKAEAAAKKAAADDK 257

QY 99 KAAAEKAAAYEA 109

DB 258 KKAATAAAKQA 268

RESULT 2

Q7RC08

ID Q7RC08 PRELIMINARY; PRT; 1701 AA.

AC Q7RC08; (TREMBlrel. 26, Created)

DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)

DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)

DE Erythrocyte binding protein.

GN Name=PY05977;

OS Plasmodium yoelii yoelii.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI\_TaxID=73239;

RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Pertea M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegh M., Shoaibi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B., O.R.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RA "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002)
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01001968; EAA18109.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
SQ SEQUENCE 1701 AA; 197910 MW; 03BE665BAE45C669 CRC64;

Query Match 37.3%; Score 193.5; DB 5; Length 1701;
Best Local Similarity 50.4%; Pred. No. 8.1e-05;
Matches 60; Conservative 12; Mismatches 28; Indels 19; Gaps 5;

QY 2 KTKAKAEKAYAKAKAEKKAY-AKKEAKYAKAEAKKAKA-----EAKYKAEAAK 55
Db 1238 KKAEAAKALERKKAEAAKALERKKAEAAKAEKKAEKKAEKKAEKKAEAAK 1297
QY 56 AKKEAYKAEKKYAKAKAEKKYAA-----BAKKAEEK-----AYKAEAAKAAKE 104
Db 1298 AEBEKKAEAA-----AKKAEBEKKAEAAKAEKKAEKKAEKKAEKKAEKKAE 1353

RESULT 3
Q9NWX1 PRELIMINARY; PRT; 372 AA.
AC Q9NWX1
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE TolA protein.
GN Name=tolA;
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RX PubMed=8626299;
RA Rodriguez-Herva J.J., Ramos-Gonzalez M.I., Ramos J.;
RT "The Pseudomonas putida peptidoglycan-associated outer membrane
RT lipoprotein (PAL) is involved in maintenance of the integrity of the
RT cell envelope."
RL J. Bacteriol. 178:1699-1706(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RX PubMed=9642202; PubMed=8824639;
RA Rodriguez-Herva J.J., Ramos J.;
RT "Characterization of an OprL null mutant of Pseudomonas putida."
RL J. Bacteriol. 178:5836-5840(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RA Ramos-Gonzalez I.;

```

```

RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RA Rodriguez-Herva J.J.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; X74218; CAB50780.1; -.
DR HSSP; P50600; 1LR0.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008565; F:nucleosome assembly; IEA.
DR GO; GO:0006334; P:nucleosome transporter activity; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR InterPro; IPR005819; Histone_H5.
DR InterPro; IPR010528; TolA.
DR InterPro; IPR006260; TonB_C.
DR Pfam; PF06519; TolA; 1.
DR PRINTS; PR00624; HISTONEH5.
DR TIGRFAMs; TIGR01352; tonB_Cterm; 1.
SQ SEQUENCE 372 AA; 40133 MW; 87F49785EC3C0BC CRC64;

Query Match 36.7%; Score 190.5; DB 2; Length 372;
Best Local Similarity 50.4%; Pred. No. 3.5e-05;
Matches 59; Conservative 14; Mismatches 31; Indels 13; Gaps 4;

QY 5 AKKAEKAYAKKAAK-----EKKAYAKKAEKAYKAE-----AKKAEKAYKAE 54
Db 118 AKKAEDA-AKAEAAKAAAEKAAKAEKAAKAEKAAKAEKAAKAEKAAKAEKAA 176
QY 55 K-AKKEAYKAEKKYAKAKAEKKYAA-----BAKKAEEK-----AYKAEAAKAAKE 109
Db 177 KKAABEAKKAAEDAKKKAABEAKKAAEDAKKKAABEAKKAAEDAKKKAABEAKK 233

RESULT 4
Q88N16 PRELIMINARY; PRT; 372 AA.
AC Q88N16
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Biopolymer transport protein TolA.
GN Name=tolA; OrderedLocustNames=P1221;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weiner C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazed A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Medler H., Lauber J., Stjepandic D., Hoesel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoef A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AB016778; AAN66845.1; -.
DR HSSP; P50600; 1LR0.
DR TIGR; P1221; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.

```

[illegible]

## RESULT 8

```
AAQ73466
ID AAQ73466 PRELIMINARY; PRT; 1830 AA.
AC AAQ73466;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-NAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Erythrocyte binding protein 2.
GN MAEBL.
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Salvador;
RX MEDLINE=22077637; PubMed=12082132;
RA Michon P., Stevens J.R., Kaneko O., Adams J.H.;
RT "Evolutionary relationships of conserved cysteine-rich motifs in
adhesive molecules of malaria parasites.";
RL Mol. Biol. Evol. 19:1128-1142(2002).
DR EMBL; AY042083; AAQ73466.1; -.
SQ SEQUENCE 1830 AA; 208690 MW; 099B2D035391E399 CRC64;

Query Match 36.7%; Score 190.5; DB 2; Length 1830;
Best Local Similarity 49.6%; Pred. No. 0.00013;
Matches 61; Conservative 17; Mismatches 22; Indels 23; Gaps 6;

QY 5 AKKAEKAYAKKAAKAE-----KKAYAKKAEKAYAKKAAKKA-----KAEAKKYA 50
Db 1262 AKKAE--ARKAEAKKAEARKAEAKKAEARKAEARKAEAKKAEARKAEARKAEARKA 1319
QY 51 KEAAKAKKAEYAKAEAKKYAKAAKAEKKEYAAAEAKKAEAAK-----AYKAEAAKAAKAA 106
Db 1320 EDARKA-EEARKAEAAKAEARKAEARKAEAKKAEAAKAEARKAEARKAEARKA--EAA 1376
QY 107 YEA 109
Db 1377 RKA 1379

RESULT 9
Q8T5C8 PRELIMINARY; PRT; 1866 AA.
AC Q8T5C8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte binding protein 1.
GN Name=maebi;
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22077637; PubMed=12082132;
RA Michon P., Stevens J.R., Kaneko O., Adams J.H.;
RT "Evolutionary relationships of conserved cysteine-rich motifs in
adhesive molecules of malaria parasites.";
RL Mol. Biol. Evol. 19:1128-1142(2002).
DR EMBL; AY042083; AAL10508.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy_binding.
DR Pfam; PF05424; Duffy_binding; 1.
SQ SEQUENCE 1866 AA; 212420 MW; DC692D7CFAE7D93F CRC64;

Query Match 36.7%; Score 190.5; DB 2; Length 1866;
Best Local Similarity 49.6%; Pred. No. 0.00014;
Matches 61; Conservative 17; Mismatches 22; Indels 23; Gaps 6;

QY 5 AKKAEKAYAKKAAKAE-----KKAYAKKAEKAYAKKAAKKA-----KAEAKKYA 50
Db 1262 AKKAE--ARKAEAKKAEARKAEAKKAEARKAEARKAEAKKAEARKAEARKAEARKA 1319

AAQ73466
ID AAQ73466 PRELIMINARY; PRT; 1830 AA.
AC AAQ73466;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-NAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Erythrocyte binding protein 2.
GN MAEBL.
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Salvador;
RX MEDLINE=22077637; PubMed=12082132;
RA Michon P., Stevens J.R., Kaneko O., Adams J.H.;
RT "Evolutionary relationships of conserved cysteine-rich motifs in
adhesive molecules of malaria parasites.";
RL Mol. Biol. Evol. 19:1128-1142(2002).
DR EMBL; AY042083; AAQ73466.1; -.
SQ SEQUENCE 1830 AA; 208690 MW; 099B2D035391E399 CRC64;

Query Match 36.7%; Score 190.5; DB 2; Length 1830;
Best Local Similarity 49.6%; Pred. No. 0.00013;
Matches 61; Conservative 17; Mismatches 22; Indels 23; Gaps 6;

QY 5 AKKAEKAYAKKAAKAE-----KKAYAKKAEKAYAKKAAKKA-----KAEAKKYA 50
Db 1262 AKKAE--ARKAEAKKAEARKAEAKKAEARKAEARKAEAKKAEARKAEARKAEARKA 1319
QY 51 KEAAKAKKAEYAKAEAKKYAKAAKAEKKEYAAAEAKKAEAAK-----AYKAEAAKAAKAA 106
Db 1320 EDARKA-EEARKAEAAKAEARKAEARKAEAKKAEAAKAEARKAEARKAEARKA--EAA 1376
QY 107 YEA 109
Db 1377 RKA 1379

RESULT 10
TOLA_ECOLI
ID TOLA_ECOLI STANDARD; PRT; 421 AA.
AC P19934;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Tola protein.
GN Name=tolA; Synonyms=cim, excC, lky; OrderedLocusNames=b0739;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / JM105;
RX MEDLINE=90078104; PubMed=2687247;
RA Levengood S.K., Webster R.E.;
RT "Nucleotide sequences of the tola and tolB genes and localization of
their products, components of a multistep translocation system in
Escherichia coli.";
RL J. Bacteriol. 171:6600-6609(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [4]
RP DOMAINS.
RX MEDLINE=91296736; PubMed=2068069;
RA Levengood S.K., Beyer W.F. Jr., Webster R.E.;
RT "Tola: a membrane protein involved in colicin uptake contains an
extended helical region.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5939-5943(1991).
RN [5]
RP INTERACTION WITH PORINS.
RX MEDLINE=97133271; PubMed=8978668;
RA Derouiche R., Gavioli M., Benedetti H., Prilipov A., Lazdunski C.,
RA Lloubes R.;
RT "Tola central domain interacts with Escherichia coli porins.";
RL EMBO J. 15:6408-6415(1996).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 298-421.
RX MEDLINE=99332679; PubMed=1040600;
RA Lubkowski J., Henneke F., Plueckthun A., Wlodawer A.;
```

Qy	56	AKKDAY-KAEAAKYKAAGAAGKEEYAAAABAKK----
Dd	178	AAAAEQKKAEAAAALKKKAEEAASARUKAATEAAEKAKAEKGAEEKAAEA
RESULT 11		
ID	Q8FUT1	PRELIMINARY; PRT; 421 AA.
AC	O8FUT1;	
DC	01-MAR-2003 (TrEMBLrel. 23, Created)	
DT	01-NAR-2003 (TrEMBLrel. 23, Last sequence update)	
DD	01-WAR-2004 (TrEMBLrel. 26, Last annotation update)	
DE	Tola protein.	
DN	Name=tola; OrderedLocusNames=c0818;	
GS	Escherichia coli O6.	
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;	
OC	Enterobacteriaceae; Escherichia.	
OX	NCBI_TaxID=217992;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=O6:H1 / CFT073 / ATCC 700928;	
RC	MEDLINE=22388234; PubMed=12471157;	
RA	Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,	
RA	Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J.J., Stroud D.	
RA	Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,	
RA	Mobley H.L.T., Donnenberg M.S., Blattner F.R.;	
RT	"Extensive mosaic structure revealed by the complete genome sequence	
RT	of uropathogenic Escherichia coli.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).	
RL	EMBL; AE016757; AACN79291.1; -.	
DR	HSSP; P19934; 1TOL.	
DR	InterPro; IPR010528; TOLA.	
DR	Pfam; PF06519; TOLA; 1.	
KW	Complete proteome.	
SQ	SEQUENCE 421 AA; 43184 MW; DB296626F056D3B5 CRC64;	
Query Match 36.3%; Score 188.5; DB 2; Length 421;		
Best Local Similarity 53.0%; Pred. No. 5.2e-05;		
Matches 61; Conservative 11; Mismatches 32; Indels 11; Gaps:		
Qy	1	AKKYAKAE-----KAYAKKAAAKEKAYAKAEKAYAKAEKAYAKAEKAYAKAEKAYAK-
Dd	120	AEEAAQKAEKQKQEAEAAAKAADAKAKAEADAKA--AEEAAKKAADAADAKKCAEAEEAA
Qy	56	AKKDAY-KAEAAKYKAAGAAGKEEYAAAABAKK----
Dd	178	AAAAEQKKAEAAAALKKKAEEAASARUKAATEAAEKAKAEKGAEEKAAEA
RESULT 12		
ID	Q74W64	PRELIMINARY; PRT; 401 AA.
AC	Q74W64;	
DC	05-JUL-2004 (TrEMBLrel. 27, Created)	
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)	
DD	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)	
DE	Tola colicin import membrane protein.	
DN	Name=tola; OrderedLocusNames=YPI033;	
GS	Yersinia pestis.	
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;	
OC	Enterobacteriaceae; Yersinia.	
OX	NCBI_TaxID=632;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=91001 / Biovar Mediaevalis;	
RC	Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.	
RA	Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,	
RA	Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,	
RA	Yang R.;	
RL	Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.	
RL	EMBL; AE017130; AAS61283.1; -.	
DR	InterPro; IPR010528; TOLA.	
DR	Pfam; PF06519; TOLA; 1.	





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 120.498 Seconds  
(without alignments)  
324.499 Million cell updates/sec

Title: US-10-792-311-7  
Perfect score: 519  
Sequence: 1 AKYAKKAEKAYAKKAKAAK.....AKAYKAEAKAAKAEAYEA 109

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	519	100.0	109	3 AAY82577	Aay82577 Copolymer
2	312.5	60.2	86	3 AAY82576	Aay82576 Copolymer
3	289	55.7	77	3 AAY82575	Aay82575 Copolymer
4	228.5	44.0	66	3 AAY82574	Aay82574 Copolymer
5	209	40.3	154	2 AAR06445	Aar06445 Recombina
6	196	37.8	469	7 ABO67048	Ab067048 Klebsiell
7	190.5	36.7	372	6 ABU40185	Abu40185 Protein e
8	190	36.6	428	6 ABU40185	Abu40185 Protein e
9	189.5	36.5	421	6 ABU28559	Abu28559 Protein e
10	187	36.0	106	2 AAR06446	Aar06446 Recombina
11	184.5	35.5	407	6 ABU47123	Abu47123 Protein e
12	181.5	35.0	323	6 ABU31397	Abu31397 Protein e
13	180.5	34.8	56	3 AAY82573	Aay82573 Copolymer
14	179.5	34.6	100	3 AAY98499	Aay98499 Peptide #
15	179.5	34.6	100	3 AAY59044	Aay59044 Amino aci
16	179.5	34.6	100	4 AAB45852	Aab45852 Nucleic a
17	179.5	34.6	100	4 AAU04289	Aau04289 Poly-Lys-
18	179	34.5	347	6 ABU38313	Abu38313 Protein e
19	179	34.5	347	6 ABJ18771	Abj18771 Pseudomon
20	179	34.5	407	7 ABO80835	Ab080835 Pseudomon
21	178	34.3	223	2 AAY14928	Aay14928 Amino aci
22	178	34.3	223	6 ABP70903	Abp70903 Mycobacte
23	178	34.3	376	6 ABU47848	Abu47848 Protein e
24	178	34.3	388	6 ABU50266	Abu50266 Protein e
25	174	33.5	387	6 ABU44721	Abu44721 Protein e

RESULT 1  
AAY82577  
ID AAY82577 standard; peptide; 109 AA.  
XX  
AC AAY82577;  
XX  
DT 28-JUL-2000 (first entry)  
XX  
DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:7.  
XX  
KW Copolymer; molecular weight marker; TV-marker; immune disease;  
KW glatiramer acetate; autoimmune disease; antirheumatic; neuroprotective;  
KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;  
KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;  
KW antianemic; immunosuppressive; demyelinating disease; osteoarthritis;  
KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;  
KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;  
KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;  
KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;  
KW pemphigus vulgaris; systemic lupus erythematosus.  
XX  
OS Unidentified.

ALIGNMENTS

26	173.5	33.4	389	6	ABU39221	Abu39221 Protein e
27	169	32.6	361	7	ADF05105	Adf05105 Bacterial
28	163.5	31.5	448	6	ABU17340	Abu17340 Protein e
29	163	31.4	336	6	ABU42038	Abu42038 Protein e
30	162.5	31.3	643	2	AAR84568	Aar84568 Trypanoso
31	162	31.2	356	6	ABU49418	Abu49418 Protein e
32	162	31.2	356	7	ABO23526	Ab023526 Vibrio ch
33	161	31.0	357	6	ABM67869	Abm67869 Photorhab
34	160.5	30.9	158	2	AAY34068	Aay34068 Histone H
35	160.5	30.9	158	2	AAY57366	Aay57366 Human his
36	160.5	30.9	226	2	AAY34060	Aay34060 Human his
37	160.5	30.9	226	2	AAY57358	Aay57358 Human his
38	160.5	30.9	226	8	ADQ09196	Adq09196 Human HLF
39	160.5	30.9	468	6	ADA35034	Ada35034 Acinetoba
40	160	30.8	214	6	ADA33882	Ada33882 Acinetoba
41	159.5	30.7	234	5	AAE13234	Aae13234 Human lin
42	159.5	30.7	234	5	AAU09944	Aau09944 Histone H
43	159.5	30.7	234	7	ADD22892	Add22892 Human his
44	158.5	30.5	472	2	AAR84569	Aar84569 Trypanoso
45	158	30.4	218	2	AAW29477	Aaw29477 Human his





```

PN EP383620-A.
XX
XX
PD 22-AUG-1990.
XX
XX
PF 16-FEB-1990; 90EP-00301700.
XX
XX
PR 17-FEB-1989; 89US-00312541.
PR 07-FEB-1990; 90US-00473845.
XX
XX
PA (REPK ) REPLIGEN CORP.
XX
XX
PI Cook KS;
XX
XX
DR WPI; 1990-255848/34.
DR N-PSDB; AAQ05664.
XX
XX
PT Producing genes encoding random polymers of aminoacid(s) - for producing
PT recombinant polypeptide(s) with biological and/or immunological activity.
XX
XX
PS Disclosure; Fig 11; 25pp; English.
XX
XX
CC To improve the expression of rCOP-1 polypeptides in E. coli, genes coding
CC for rCOP-1-77 were subcloned from pREV 2.1 to pBG3-2deltaN (deposit: 20-
CC NOV-1984 US4691009, NHRB B-15910), a plasmid used to express Protein A.
CC The resulting plasmids encode fusion proteins consisting of beta-
CC glucuronidase, Protein A, and rCOP-1 sequences. A methionine residue
CC occurs between the Protein A and rCOP-1 sequences, originating from the
CC 5' linker sequence, in order that the COP-1 polypeptide may be cleaved
CC from the fusion protein. rCOP-1-77 contains oligonucleotide duplexes
CC encoding the following segments: YKK, EAE, KAK, AAK, and AAA. The N-
CC terminal alanine residue is left behind following CNBr cleavage of the
CC fusion protein. The product prevents or arrests experimental autoimmune
CC encephalomyelitis. They are used to prevent, arrest or control a
CC demyelinating disorder, e.g. multiple sclerosis. They may also be used as
CC additives to hair care products to confer beneficial effects on damaged
CC hair or as supplements for diets deficient in certain amino acids. See
CC also AAQ05665. (Updated on 25-MAR-2003 to correct PA field.)
XX
XX
SQ Sequence 154 AA;

Query Match 40.3%; Score 209; DB 2; Length 154;
Best Local Similarity 53.7%; Pred. No. 3.9e-10;
Matches 72; Conservative 8; Mismatches 20; Indels 34; Gaps 9;

QY 2 KKYAKAEKA-----YAKKAAAEKKKAYA---KKEAKAYKAEAE-----KKAK-A 44
DB |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
24 KKAKEAEKAKKAKYKKEAEAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 81
QY 45 EAKKYAKEAKA--KKEAYKAEAKKYAKAKAEK-----KEYAAAEAKKAEAA--K 91
DB |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
82 EYKKAKAAAEAYKKEAEAYKKEAYKKEAYKKEAYKKEAYKKEAYKKEAYKKEAYK 141
QY 92 AYKAEAAKAAKAE 105
DB |||:||||| |||:|||||
142 KYKKEAEK--AKEA 153

RESULT 6
ABO67048
ID ABO67048 standard; protein; 469 AA.
XX
XX
AC ABO67048;
XX
XX
DT 29-JUL-2004 (first entry)
XX
DE Klebsiella pneumoniae polypeptide seqid 13565.
XX
KW Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX
XX
OS Klebsiella pneumoniae.
XX
XX
PN US6610836-B1.

```

```

XX
PD 26-AUG-2003.
XX
XX
PF 27-JAN-2000; 2000US-00489039.
XX
XX
PR 29-JAN-1999; 99US-0117747P.
XX
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
XX
PI Breton GL, Osborne M;
XX
XX
DR WPI; 2003-895346/82.
DR N-PSDB; ABD00619.
XX
XX
PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
XX
XX
PS Disclosure; SEQ ID NO 13565; 932pp; English.
XX
XX
CC The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This is the amino acid sequence of a
CC Klebsiella pneumoniae polypeptide of the invention
XX
XX
SQ Sequence 469 AA;

Query Match 37.8%; Score 196; DB 7; Length 469;
Best Local Similarity 51.2%; Pred. No. 1.5e-08;
Matches 62; Conservative 12; Mismatches 35; Indels 12; Gaps 4;

QY 1 AKKYAKAEKAYKAKAEK---KAYAKKAEKAYKAEKAKKAEKAKKAEKAKKAEKAKKAEK 53
DB |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
153 AKKQKAEAEAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 212
QY 54 AKAKKEAYKAEKAYKAKA---AKAEKKEVAAAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 108
DB |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
213 ADAKQAEAEAAKAAAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 272
QY 109 A 109
DB 273 A 273

RESULT 7
ABU40185
ID ABU40185 standard; protein; 372 AA.
XX
XX
AC ABU40185;
XX
XX
DT 19-JUN-2003 (first entry)
XX
XX
DE Protein encoded by Prokaryotic essential gene #25712.
XX
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX
OS Pseudomonas putida.
XX
XX
PN WO200277183-A2.
XX
XX
PD 03-OCT-2002.
XX
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX
PA (ELIT-) ELITRA PHARM INC.

```















***This Page Blank (uspto)***

Run on: December 14, 2004, 05:52:22 ; Search time 270.43 Seconds  
(without alignments)  
143.965 Million cell updates/

Sequence: 1 AKKYAKKAEKAYAKKAKAAK.....AKAYKAEAAKAAAKEAAAYEA 109

Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 15855576

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%

```
Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/FCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/FCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query			Description	
	Score	Match	Length	ID	
1	519	100.0	109	9	US-09-816-989A-7
2	312.5	60.2	86	9	US-09-816-989A-6
3	289	55.7	77	9	US-09-816-989A-5
4	228.5	44.0	66	9	US-09-816-989A-4
5	190.5	36.7	372	15	US-10-282-122A-68109
6	190	36.6	428	15	US-10-282-122A-55748
7	189.5	36.5	421	15	US-10-282-122A-56483
8	184.5	35.5	407	15	US-10-282-122A-75047
9	181.5	35.0	323	15	US-10-282-122A-59321
10	180.5	34.8	56	9	US-09-816-989A-3
11	179	34.5	347	14	US-10-127-032-120
12	179	34.5	347	15	US-10-282-122A-66237
13	178	34.3	223	13	US-10-051-643-201

Db 61 YKAEAKKYAKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 109

RESULT 2

US-09-816-989A-6

; Sequence 6, Application US/09816989A

; Patent No. US20020115103A1

; GENERAL INFORMATION:

; APPLICANT: Gad, Alexander

; APPLICANT: Lis, Doris

; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK

; FILE REFERENCE: 2609/60807-A-PCT-US

; CURRENT APPLICATION NUMBER: US/09/816,989A

; CURRENT FILING DATE: 2001-03-23

; PRIOR APPLICATION NUMBER: 60/101,693

; PRIOR FILING DATE: 1998-09-25

; PRIOR APPLICATION NUMBER: PCT/US99/22402

; PRIOR FILING DATE: 1999-09-24

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 6

; LENGTH: 86

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide

US-09-816-989A-6

Query Match 60.2%; Score 312.5; DB 9; Length 86;

Best Local Similarity 72.1%; Pred. No. 1.3e-17;

Matches 80; Conservative 1; Mismatches 3; Indels 27; Gaps 5;

Qy 1 AKKYAKKAEKAYAKKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 60

Db 1 AKKYAKK-EKAYAKKA-----EKAACKAEKAYKAAAEAKKKA----- 36

Qy 61 YKAEAKKYAKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 109

Db 37 -KAEAKKYAKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 86

RESULT 3

US-09-816-989A-5

; Sequence 5, Application US/09816989A

; Patent No. US20020115103A1

; GENERAL INFORMATION:

; APPLICANT: Gad, Alexander

; APPLICANT: Lis, Doris

; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK

; FILE REFERENCE: 2609/60807-A-PCT-US

; CURRENT APPLICATION NUMBER: US/09/816,989A

; CURRENT FILING DATE: 2001-03-23

; PRIOR APPLICATION NUMBER: 60/101,693

; PRIOR FILING DATE: 1998-09-25

; PRIOR APPLICATION NUMBER: PCT/US99/22402

; PRIOR FILING DATE: 1999-09-24

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 5

; LENGTH: 77

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide

US-09-816-989A-5

Query Match 55.7%; Score 289; DB 9; Length 77;

Best Local Similarity 67.0%; Pred. No. 7.9e-16;

Matches 73; Conservative 1; Mismatches 3; Indels 32; Gaps 4;

Qy 1 AKKYAKKAEKAYAKKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 60

Db 1 AKKYAKK-EKAYAKKA-----EKAACKAEKAYKAAAEAKKKA----- 36

Qy 61 YKAEAKKYAKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 109

Db 37 -KAEAKKYAKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 77

RESULT 4

US-09-816-989A-4

; Sequence 4, Application US/09816989A

; Patent No. US20020115103A1

; GENERAL INFORMATION:

; APPLICANT: Gad, Alexander

; APPLICANT: Lis, Doris

; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK

; FILE REFERENCE: 2609/60807-A-PCT-US

; CURRENT APPLICATION NUMBER: US/09/816,989A

; CURRENT FILING DATE: 2001-03-23

; PRIOR APPLICATION NUMBER: 60/101,693

; PRIOR FILING DATE: 1998-09-25

; PRIOR APPLICATION NUMBER: PCT/US99/22402

; PRIOR FILING DATE: 1999-09-24

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 4

; LENGTH: 66

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide

US-09-816-989A-4

Query Match 44.0%; Score 228.5; DB 9; Length 66;

Best Local Similarity 56.9%; Pred. No. 3.6e-11;

Matches 62; Conservative 0; Mismatches 4; Indels 43; Gaps 4;

Qy 1 AKKYAKKAEKAYAKKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 60

Db 1 AKKYAKK-EKAYAKKA-----EKAACKKA----- 25

Qy 61 YKAEAKKYAKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 109

Db 26 -KAEAKKYAKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 66

RESULT 5

US-10-282-122A-68109

; Sequence 68109, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

```
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 68109
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-282-122A-68109

Query Match      36.7%; Score 190.5; DB 15; Length 372;
Best Local Similarity 50.4%; Pred. No. 2e-07;
Matches 59; Conservative 14; Mismatches 31; Indels 13; Gaps 4;

QY 5 AKKAEKAYAKKAAK---EKKAYAKKAEKAYKAAE-----AKKAKAEAKKAYKAAE 54
Db 118 AKKAEDA-AKAAEAAKAAKAAKAAEAKKAADEAKKAAEKQOADIJAKKAAEAEAKKAAEBA 176

QY 55 K-AKKEAYKAEAKKAYAKKAAKAEKKEYYAAAEAKKAEAAKAYKAAEAAKAAEAAEBA 109
Db 177 KKAABEAAKKAEDAKKAAEBAKKAEDAKKAAEDAKKAAEBAKKAEDAKKAAEBAKKAADA 233

RESULT 6
US-10-282-122A-55748
; Sequence 55748, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-11-27
; Remaining Prior Application data removed - See File Wrapper or PALM.

; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 55748
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55748

Query Match      36.6%; Score 190; DB 15; Length 428;
Best Local Similarity 50.8%; Pred. No. 2.5e-07;
Matches 64; Conservative 8; Mismatches 32; Indels 22; Gaps 5;

QY 1 AKKYA-----KKAEKAYAKKAAKAEKAYKAAEAKKAYKAAEAKKAAEAE-----K 47
Db 167 AKKAADAAQKKAEEAAKKAADAAQKKAEEAAKKA--AADQKAAEAAKKAQAEAK 224

QY 48 KYAKEAAKAKKAEYKAEAKKAYKAAKAAKAEKKEYYAAAE-----KKAEEAAKA 100
Db 225 KAAAEAAKKAABEAKKAAEAKKAAEAKKAAADKKAABEAKKAAADKKAADKKAADKA 284

QY 101 AAKEAA 106
Db 285 AAKKAA 290

RESULT 7
US-10-282-122A-56483
; Sequence 56483, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-11-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

[illegible]







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 30.5844 Seconds  
(without alignments)  
236.351 Million cell updates/sec

Title: US-10-792-311-7  
Perfect score: 519  
Sequence: 1 AKKYAKAEKAYAKKAKAAK.....AKAYKAEAAKAAAKEAAAYEA 109

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

**Total number of hits satisfying chosen parameters: 478139**

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```

Database :
1 : /cgn2_6/ptodata/1/iaa/5A_COMB pep.*
2 : /cgn2_6/ptodata/1/iaa/5B_COMB pep.*
3 : /cgn2_6/ptodata/1/iaa/6A_COMB pep.*
4 : /cgn2_6/ptodata/1/iaa/6B_COMB pep.*
5 : /cgn2_6/ptodata/1/iaa/PCTUS_COMB pep.*
6 : /cgn2_6/ptodata/1/iaa/backfiles1 pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	519	100.0	109	4	US-09-405-743A-7	Sequence 7, Appli
2	312.5	60.2	86	4	US-09-405-743A-6	Sequence 6, Appli
3	289	55.7	77	4	US-09-405-743A-5	Sequence 5, Appli
4	228.5	44.0	66	4	US-09-405-743A-4	Sequence 4, Appli
5	196	37.8	469	4	US-09-489-039A-13565	Sequence 13565, A
6	180.5	34.8	56	4	US-09-405-743A-3	Sequence 3, Appli
7	179.5	34.6	100	2	US-08-460-890A-64	Sequence 64, Appl
8	179.5	34.6	100	3	US-08-167-641C-64	Sequence 64, Appl
9	179.5	34.6	100	3	US-08-460-971A-64	Sequence 64, Appl
10	179.5	34.6	100	3	US-08-462-040-64	Sequence 64, Appl
11	179	34.5	407	4	US-09-252-991A-29581	Sequence 29581, A
12	178	34.3	223	3	US-09-095-855-201	Sequence 201, App
13	178	34.3	223	4	US-09-205-426-201	Sequence 201, App
14	169	32.6	361	4	US-09-543-681A-5390	Sequence 5390, Ap
15	164	31.6	171	4	US-09-270-767-39148	Sequence 39148, A
16	164	31.6	171	4	US-09-270-767-39148	Sequence 39148, A
17	162.5	31.3	643	2	US-08-216-894-8	Sequence 54365, A
18	162.5	31.3	643	3	US-09-115-746-8	Sequence 8, Appli
19	160.5	30.9	158	3	US-09-041-889-40	Sequence 40, Appl
20	160.5	30.9	158	4	US-09-417-264-40	Sequence 40, Appl
21	160.5	30.9	226	3	US-09-041-889-32	Sequence 32, Appl
22	160.5	30.9	226	4	US-09-417-264-32	Sequence 32, Appl
23	160.5	30.9	468	4	US-09-328-352-6321	Sequence 6321, Ap
24	160	30.8	214	4	US-09-328-352-6321	Sequence 6321, Ap
25	158.5	30.5	472	2	US-08-216-894-10	Sequence 10, Appl
26	158.5	30.5	472	3	US-09-115-746-10	Sequence 10, Appl
27	158.5	30.5	564	2	US-08-216-894-2	Sequence 2, Appli

28	158.5	30.5	564	3	US-09-115-746-2	Sequence 2, Appli
29	158	30.4	218	3	US-09-041-889-4	Sequence 4, Appli
30	158	30.4	218	3	US-08-837-050-4	Sequence 4, Appli
31	158	30.4	218	4	US-09-417-264-4	Sequence 4, Appli
32	152.5	29.4	222	3	US-09-041-889-3	Sequence 3, Appli
33	152.5	29.4	222	3	US-08-837-058-3	Sequence 3, Appli
34	152.5	29.4	222	4	US-09-417-264-3	Sequence 3, Appli
35	151	29.1	316	4	US-09-259-991A-32957	Sequence 32957, A
36	150	28.9	1507	3	US-08-929-329-5	Sequence 5, Appli
37	148	28.5	207	4	US-09-489-039A-13743	Sequence 13743, A
38	144	27.7	399	4	US-09-252-981A-22853	Sequence 22853, A
39	144	27.7	434	1	US-08-097-830E-3	Sequence 3, Appli
40	144	27.7	434	2	US-08-456-112B-3	Sequence 3, Appli
41	143.5	27.6	921	4	US-09-543-681A-5734	Sequence 5734, Ap
42	143	27.6	214	3	US-09-041-889-27	Sequence 27, Appl
43	143	27.6	214	4	US-09-417-264-27	Sequence 27, Appl
44	141	27.2	100	2	US-08-460-890A-62	Sequence 62, Appl
45	141	27.2	100	3	US-08-167-641C-62	Sequence 62, Appl

## ALIGNMENTS

```

RESULT 1
US-09-405-743A-7
; Sequence 7, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co., Ltd.
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
; FILE REFERENCE: 60807-A
; CURRENT APPLICATION NUMBER: US/09/405,743A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 7
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: PEPTIDE
US-09-405-743A-7
SYNTHETIC

```

	Query Match	100.0%;	Score 519;	DB 4;	Length 109;
	Best Local Similarity	100.0%;	Pred. No. 7.1e-37;		
	Matches 109;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AKKYAKAEKAYAKAAKEKKYAYAKKEKAYAAAEAKKKAABAKKYAKAEAAKKEA	60		
Db	1	AKKYAKAEKAYAKAAKAAKEKKYAYAKKEKAYAAAEAKKKAABAKKYAKAEAAKKEA	60		
Qy	61	YKAEEAKKYAKAAKAEKKKYAAAABAKKAAEAAKAYAAEAKAAAEKAAAYEA	109		
Db	61	YKAFAKKYAKAAKAEKKKYAAAABAKKAAEAAKAYAAEAKAAAEKAAAYEA	109		

RESULT 2  
US-09-405-743A-6  
; Sequence 6, Application US/09405743A  
; Patent No. 6514938  
; GENERAL INFORMATION:  
; APPLICANT: Yeda Research and Development Co., Ltd.  
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS  
; FILE REFERENCE: 60807-A  
; CURRENT APPLICATION NUMBER: US/09/405, 743A  
; CURRENT FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.1

```

; TYPE: PRI
; ORGANISM: Artificial Sequence

```







```

; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29581
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29581

Query Match 34.5%; Score 179; DB 4; Length 407;
Best Local Similarity 49.1%; Pred. No. 6.7e-08;
Matches 54; Conservative 17; Mismatches 29; Indels 10; Gaps 4;

Qy 6 KKA EKAYAKKAKAAK-----KAYAKKEAKYKAAEAKKKAKAEAKKYAKAEAAKKE 59
Db 159 QKLEQQVAAAKAAEQKKADEAKFAEQKAAEAKGADAEAKKAAEAKAAEQKQADIAKGR 218
Qy 60 AYKAEAKKYAKAAKAEKKEYYAAAEAKAEAAKAYKAEAAKAAAKAEAAEAAVEA 109
Db 219 A-EDEAKK--KAAEDAKKK-AAEDAKKAAEAEAKKAAAEAAEAAKKAAVEA 264

RESULT 12
US-09-095-855-201
; Sequence 201, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
; TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,855
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; APPLICATION NUMBER: 08/873,970
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: 08/997,362
; FILING DATE: 23-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565

```

```
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5390
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5390

Query Match      32.6%; Score 169; DB 4; Length 361;
Best Local Similarity 44.0%; Pred. No. 4e-07;
Matches 51; Conservative 16; Mismatches 37; Indels 12; Gaps 2;

QY 1 AKKYAKAEKAYAKKAAKAEKAYAKKAEKAYAKA-----AEAKKAKAEAKKAYAK 51
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 123 AKEAAEQRRREATAAAKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 182
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 52 EAAKAKAEKAYAKKAEKAYAKA---AKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 104
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 183 LAAKQAEAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEK 238

RESULT 15
US-09-270-767-39148
; Sequence 39148, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39148
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-39148

Query Match      31.6%; Score 164; DB 4; Length 171;
Best Local Similarity 42.9%; Pred. No. 5e-07;
Matches 57; Conservative 20; Mismatches 30; Indels 26; Gaps 6;

QY 2 KKYAKAEKAYAKKAAKAEKK-----AYAKKE--AKAYKAEAKKK-----AKAEAK 47
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 22 KKAQKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 81
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 48 KYAKAEA-----KAKKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 96
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 82 KAAEAAKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 140
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 97 AAKAAAEKAEAYEA 109
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 141 AAQKAEAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 153
```

Search completed: December 14, 2004, 05:50:19  
Job time : 31.7272 secs